thematics

C-27 NOV 1 1937

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

J. F. RITT
COLUMBIA UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY	MARSTON MORSE	G. C. EVANS
W. A. MANNING	E. P. LANE	AUREL WINTNER
HARRY BATEMAN	ALONZO CHURCH	GABRIEL SZEGŐ
HARRY LEVY	L. R. FORD	R. L. WILDER
J. R. KLINE	OSCAR ZARISKI	R. D. JAMES

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LIX, Number 4

OCTOBER, 1937

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

	PAGE
On Dirichlet's problem. By C. Carathéodory,	709
Bounded analytic functions in several variables and multiple Laplace integrals. By S. Bochner,	732
The characters of the symmetric group. By F. D. Murnaghan,	739
Isotropic static solutions of the field equations in Einstein's theory of gravitation. By P. Y. Chou,	754
On tensors relative to the extended point transformation. By H. V. Craig,	764
A series of involutorial Cremona transformations in S_n belonging multiply to a non-linear line complex. By Virgil Snyder and Evelyn Carroll-Rusk,	775
The Veneroni transformation in S_n. By Gertrude Blanch,	783
On $2n$ points with a real cross-ratio. By F. Morley and J. R. Musselman,	787
Extract from a letter by E. Cartan concerning my note: On closed spaces of constant mean curvature. By T. Y. Thomas,	793
Almost periodic functions and Hill's theory of lunar perigee. By Aurel Wintner,	795
Sur les mouvements des systèmes dynamiques qui admettent "l'incompressibilité" des domaines. Par Heinrich Hilmy,	803
The structure of monotone functions. By Philip Hartman and Richard Kershner,	809
Planar graphs whose homeomorphisms can all be extended for any mapping on the sphere. By V. W. Adkisson and Saunders MacLane,	823
Geometric characterizations of invariant partial differential equations. By Aaron Fialkow,	833
Sets of conjugate matrices. By E. T. Browne,	845
Abelian fields and duality of abelian groups. By Reinhold Baer,	869
A note on topological fields. By N. Jacobson,	889
The quaternionic congruence $\lambda at \equiv b \pmod{g}$ and the equation $h(8n+1) = x_1^2 + x_2^2 + x_3^2$. By Gordon Pall,	895
A list of expansions for the function $\phi(x, y, z)$. By Walter H. Gage,	914
Some arithmetical applications of residuation. By Morgan Ward,	921
Rational curves of order $n+2$ invariant under dihedral collineation groups of order $2n$. By R. M. Winger,	927
Contribution à l'étude des systèmes de choses normées. Par V. Glivenko,	941
Topological properties of the solutions of ordinary differential equations. By T. M. Cherry,	957
Theory of quasi-groups. By B. A. Hausmann and Oystein Ore,	983

The American Journal of Mathematics will appear four times yearly.

The subscription price of the Journal for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the Journal remain on sale.

Papers intended for publication in the Journal may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. Cohen at The Johns Hopkins University.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore, Maryland, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

ON DIRICHLET'S PROBLEM.*

By C. CARATHÉODORY.

During the last 75 years Dirichlet's Problem has attracted the attention of many of the best scholars of the time with the result that this problem, which once was considered as one of the most difficult subjects in higher mathematics, can be treated now in a very elementary way.

Carl Neumann and *H. A. Schwarz* were the first to consider the problem of Dirichlet with modern methods, but while the road laid out by *C. Neumann* leads to the theory of integral equations, it is the work of *Schwarz* which has influenced most of the subsequent proofs for Dirichlet's problem. It contains already the fundamental idea of building up by a convergent iterative process very general classes of solutions by using the particular solutions given by Poisson's integral for the sphere.

The "sweeping out" process devised fifteen or twenty years later by *Poincaré* lies in the same line of thought but represents an immense advance if one compares it with the early method which *Schwarz* had developed chiefly for the case of the logarithmic potential and for very special boundaries. The splendid way in which *Poincaré*, taking advantage of the progresses made in the meantime by the theory of point sets, deals with the problem, has given many hints to younger scientists and has contributed to many advances made subsequently in different fields of mathematical science. But originally the method of *Poincaré* was very complicated and cumbersome; he felt, in fact, obliged to split the functions to which he was applying his process into two parts for which his convergence proof was holding, and to do this he had to approximate these functions by polynomials. Thus many elements foreign to potential theory were implied in his proof and the feeling that this proof was likely to be simplified and made clearer in its outline was present from the beginning to many students of the topic.

But it was not until 1912 that a most decisive and very surprising turn was given to the whole subject in a short note of 3 pages by *H. Lebesgue*. In this note it was shown that Poisson's Integral, which up to that time had been the primordial tool with which the convergent processes were built up, could be discarded altogether and replaced by much weaker operators connected with the property of the mean which characterizes harmonic functions. These

* Received February 28, 1937.

results of *Lebesgue*, though widely quoted, are not as universally known as would have been the case if their importance had been fully appreciated from the outset.

The great variety of convergent processes outlined above has induced me to look at the common root out of which all these processes originate. The result is contained in the second chapter of this memoir. It consists of a theorem whose assumptions are taken general enough so as to disclose the very reasons for which all the different methods used previously are equally successful.

In the foregoing chapter I have tried to give a very elementary treatment of the principal properties of harmonic functions culminating in the existence proof for Dirichlet's problem devised by *O. Perron* and very much simplified by *T. Radò* and *F. Riesz*. I have done this in order to show how the whole theory can be condensed if one puts systematically from the outset Poisson's Integral in the limelight.

Chapter I. Existence of the Solution of Dirichlet's Problem.

1. Definition of harmonic functions. Functions $U(x_1, \cdots, x_n)$, which are defined in some region R of the Euclidean space with the coördinates

$$x_1, \cdots, x_i, \cdots x_n$$

are called harmonic in that region if they are first of all continuous, if secondly considering them as functions of every single variable x_i alone they have continuous first derivatives $\partial U/\partial x_i$ and finite second derivatives $\partial^2 U/\partial x_i^2$ and if the expression

$$\Delta U = \frac{\partial^2 U}{\partial x_1^2} + \cdots + \frac{\partial^2 U}{\partial x_n^2} \tag{1.1}$$

formed with these second derivatives vanishes identically throughout R.

We shall see later on (§ 6) that every harmonic function is analytic and that consequently all of the successive partial derivatives exist and are continuous in all the variables. But to prove this we do not need to assume the existence of other partial derivatives than those which we have mentioned above or to assume that these derivatives are bounded or that the first derivatives $\partial U/\partial x_i$ are continuous functions of other variables than that which appears in the denominator.

In order to study under these assumptions the most general solutions of the equation

$$\Delta U = 0 \tag{1.2}$$

we must first obtain some special solutions of that equation.

Taking a point x_i^0 in our space and putting

$$r^2 = (x_1 - x_1^0)^2 + \cdots + (x_n - x_n^0)^2 \tag{1.3}$$

we have for every function of the form $F(r)$ the identity

$$\Delta F = F''(r) + F'(r)\,\frac{n-1}{r}. \tag{1.4}$$

The most general harmonic functions of the form $F(r)$ are therefore given by the equation

$$F(r) = \frac{A}{r^{n-2}} + B \tag{1.5}$$

for $n \geqq 3$ and by

$$F(r) = A \log r + B \tag{1.6}$$

for $n = 2$.

2. Poisson's integral. Taking first $n \geqq 3$ we consider a sphere of radius a and on some radius of this sphere two points at the distances $\lambda(\lambda < a)$ and a^2/λ from the center respectively. It follows then easily by the consideration of similar triangles that for points Q taken anywhere on the surface of the sphere one has

$$\frac{r'_0}{r_0} = \frac{a}{\lambda}$$

and herefrom that the function

$$V_\lambda = \sigma\left[\left(\frac{1}{r}\right)^{n-2} - \left(\frac{a}{\lambda r'}\right)^{n-2}\right] \tag{2.1}$$

is a harmonic function which vanishes identically at the surface of the sphere. We have now from the triangles of the figure on page 712

$$\frac{\lambda^2}{a^2}\,r'^2 = \frac{\lambda^2}{a^2}\left[\frac{a^4}{\lambda^2} + \rho^2 - 2\,\frac{a^2}{\lambda}\,\rho\cos\theta\right]$$

$$= a^2 + \frac{\lambda^2}{a^2}\,\rho^2 - 2\lambda\rho\cos\theta,$$

$$r^2 = \lambda^2 + \rho^2 - 2\lambda\rho\cos\theta,$$

so that we can write

$$\frac{\lambda^2}{a^2}\,r'^2 = r^2 + u \tag{2.2}$$

with

$$u = \frac{(a^2 - \lambda^2)(a^2 - \rho^2)}{a^2}. \tag{2.3}$$

By these formulas we can write

$$\left(\frac{a}{\lambda r'}\right)^{n-2} = (r^2 + u)^{-[(n-2)/2]}; \tag{2.4}$$

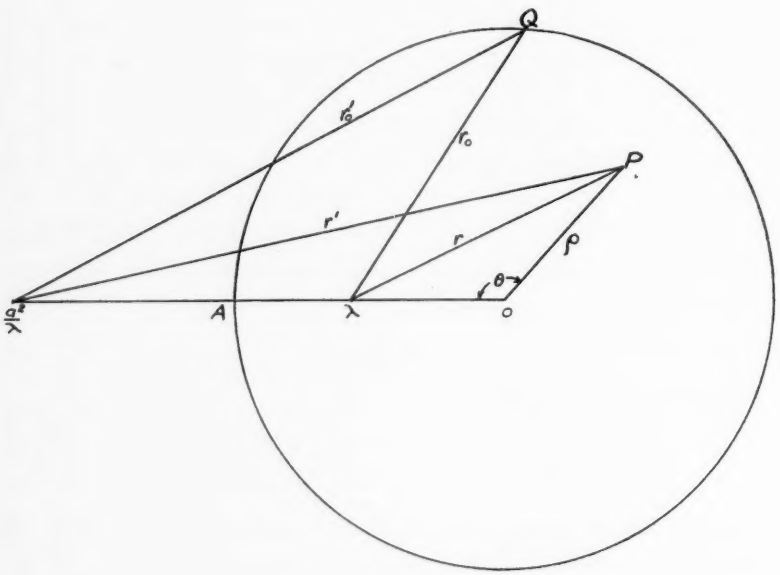
therefore if we put

$$\omega = a^2 - \lambda^2 \tag{2.5}$$

and, observing that r depends not only upon the x_i but also upon λ,

$$\left(r^2 + \frac{(a^2 - \rho^2)\omega}{a^2}\right)^{-[(n-2)/2]} = \phi(x_i, \lambda, \omega), \tag{2.6}$$

we can calculate the right hand of (2.4) by $\phi(x_i, \sqrt{a^2 - \omega}, \omega)$ for the value of ω given by (2.5). With these notations we can therefore replace (2.1) by



$$V_\lambda = -\sigma\omega \frac{\phi(x_i, \sqrt{a^2 - \omega}, \omega) - \phi(x_i, \sqrt{a^2 - \omega}, 0)}{\omega}. \tag{2.7}$$

We shall now take $\sigma\omega$ to be a constant for varying ω and make use of the fact that the function

$$\psi(x_i, \lambda, \omega) = \frac{\phi(x_i, \lambda, \omega) - \phi(x_i, \lambda, 0)}{\omega} \tag{2.8}$$

is analytic if the x_i vary in a suitable domain and that, $|a - \lambda|$ and $|\omega|$ being sufficiently small, the point $\omega = 0$ is not excluded from the domain of regularity of that function. We have in fact

$$\psi(x_i, \lambda, 0) = \frac{\partial \phi}{\partial \omega}\Big|_{\omega = 0} = -\frac{n-2}{2}\,\frac{1}{r^n}\,\frac{a^2 - \rho^2}{a^2}. \tag{2.9}$$

Taking therefore

$$\sigma\omega = C\frac{2a^2}{n-2} \tag{2.10}$$

the function V_λ takes at the limit $\omega = 0$ or, which is the same, at the limit $\lambda = a$ the value

$$V_a = \sigma\omega\psi(x_i, a, 0) = C\frac{a^2-\rho^2}{r^n}. \tag{2.11}$$

Now we have for real and positive values of ω

$$\Delta\psi(x_i, \sqrt{a^2-\omega}, \omega) = -\frac{n-2}{2a^2C}\Delta V_\lambda = 0$$

and the analytic function $\Delta\psi(x_i, \sqrt{a^2-\omega}, \omega)$ being regular for $\omega = 0$ we see that the right hand of (2.11) is *harmonic* at every real point of the space except at the point A of the above figure.

3. In order to show that the formula (2.11) holds also for the case of the two dimensional plane we observe that the function

$$V_\lambda = \sigma\left[\log\frac{\lambda}{a}r' - \log r\right] \tag{3.1}$$

is harmonic and vanishes at the boundary of the circle with radius a and that by using our previous notations we can write

$$\begin{aligned} V_\lambda &= \frac{\sigma}{2}\left[\log(r^2+u) - \log r^2\right] \\ &= \frac{\sigma}{2}\log\left(1+\frac{a^2-\rho^2}{a^2r^2}\omega\right) \end{aligned} \tag{3.2}$$

Putting now

$$\phi(x_i, \lambda, \omega) = \log\left(1+\frac{a^2-\rho^2}{a^2r^2}\omega\right) \tag{3.3}$$

we find that (2.7) holds also in the case $n = 2$ and that we get by the same method as above the relation (2.11), if we choose σ to satisfy the relation

$$\sigma\omega = C\cdot 2a^2. \tag{3.4}$$

4. We consider now a continuous function $V(Q)$ on the boundary of the sphere (or the circle) of radius a and form the integral

$$V(P) = C\int V(Q)\frac{a^2-\rho^2}{r^n}\,do \tag{4.1}$$

in which P is a point of the interior of the considered domain at the distance ρ from its center and r designates the distance between P and Q. One verifies

by differentiation that the function $V(P)$ is harmonic in the interior of the sphere. It is easy to calculate this function for the case when $V(Q)$ is a constant. In fact $V(P)$ is then a harmonic function of the form $F(\rho)$ which is bounded in the neighborhood of the center of the sphere and which by the formulas (1. 5) and (1. 6) must, therefore, in any case be a *constant*. At the center of the sphere we have $\rho = 0$ and $r \equiv a$ and therefore

$$V(P) = V(Q) \frac{C}{a^{n-2}} \int do.$$

This last integral has the value $\alpha_n a^{n-1}$ where α_n is a known constant depending only on the dimension n of the space. In order to have $V(P) \equiv V(Q)$ we must therefore take

$$C = \frac{1}{\alpha_n a} \tag{4. 3}$$

and get the final formula

$$V(P) = \frac{1}{\alpha_n a} \int V(Q) \frac{a^2 - \rho^2}{r^n} do \tag{4. 4}$$

which is known by the name of *Poisson's integral.*

5. From the last reasonings we infer that the identity holds

$$\frac{1}{\alpha_n a} \int \frac{a^2 - \rho^2}{r^n} do \equiv 1. \tag{5. 1}$$

With the aid of this identity *H. A. Schwarz* has proved that the function $V(P)$ represented by the relation (4. 4) is continuous on the *closed* sphere and that the boundary values of $V(P)$ coincide with the given values of $V(Q)$.

We omit the proof of this theorem of *Schwarz* which is classical and is contained in all the books of reference dealing with our subject.

6. The fundamental theorem. The proposition from which we can deduce nearly all the principal properties of harmonic functions consists in the statement that in the neighborhood of a point we may always represent any given harmonic function by a suitably chosen Poisson's integral.

We consider a function $f(P)$ which is continuous on a closed sphere and harmonic in the interior of it according to the definition of § 1, and calculate the function $V(P)$ which is given by the integral of Poisson and which possesses the same boundary values as $f(P)$. Furthermore we calculate by the integral of Poisson a function $\psi(P)$ which is harmonic in the interior of the sphere and takes on its surface the same boundary values as the function

$$\tfrac{1}{2}x_1^2.$$

Putting now

$$\Phi(P) = (f(P) - V(P)) + \alpha(\psi(P) - \tfrac{1}{2}x_1^2) \tag{6.1}$$

we observe that the function $\Phi(P)$ is continuous on the closed sphere, vanishes at its boundary and that the equality

$$\Delta\Phi + \alpha = 0 \tag{6.2}$$

holds at all the interior points of the sphere. From this we infer that for positives values of the parameter α

$$\Phi \geqq 0 \tag{6.3}$$

everywhere in the sphere. For if at some point of the sphere Φ were negative there should exist at least one point P_0 of the interior of the sphere at which $\Phi(P)$ would attain its minimum value. At such a point we should have necessarily

$$\frac{\partial\Phi}{\partial x_i} = 0, \qquad \frac{\partial^2\Phi}{\partial x_i^2} \geqq 0 \qquad (i = 1, 2, \cdots, n), \tag{6.4}$$

i. e., conditions which contradict (6.2). Letting now α tend towards zero we find that we must have everywhere in the sphere $V(P) \leqq f(P)$. We prove in the same way the inequality $-V(P) \leqq -f(P)$ so we get finally

$$V(P) \equiv f(P) \tag{6.5}$$

and our fundamental theorem is thus proved.

7. The theorem of Gauss and its converse. In order to calculate Poisson's integral at the center O of the sphere we have to put in (4.4)

$$\rho = 0, \qquad r \equiv a$$

and we get

$$V(O) = \frac{1}{\alpha_n a^{n-1}} \int V(Q)\, do. \tag{7.1}$$

This formula shows that the value of $V(P)$ at the center O of our sphere is equal to the mean of the values $V(Q)$ at the surface of that figure.

Considering the fundamental theorem of the previous section we see that this property applies to the most general harmonic functions and to every sphere which lies entirely inside of the region for which the function is defined. This statement is called the theorem of Gauss.

8. Take now a continuous function $g(P)$ for which the theorem of Gauss holds for every sphere lying inside of a region R. Suppose that $g(P)$ is not a

constant and call M and m the upper and the lower bound of $g(P)$ when P is describing the region. We are going to show that at every point P_0 of R we always have

$$m < g(P_0) < M \tag{8.1}$$

the equality signs being excluded. Call R' that part of R at each point Q of which we have

$$g(Q) \neq g(P_0). \tag{8.2}$$

The point set R' is an open set which ($g(P)$ not being constant) is not void and does not fill out the region R. There exist therefore at least one boundary point P_1 of R' lying inside of R and at this point we have

$$g(P_1) = g(P_0). \tag{8.3}$$

Furthermore there exists within R at least a sphere with center P_1 on whose surface there is at least one point Q of R'. The function $g(P)$ cannot be constant throughout the surface of this sphere for in that case we should have by Gauss' theorem $g(P_1) = g(Q)$ and this is contrary to our hypothesis. The function $g(P)$ being continuous and not constant on the surface of the sphere the mean $g(P_1)$ of these values must actually lie between M and m and by (8.3) this is equivalent to the relation (8.1) we wanted to prove.

9. From the last result it follows that if a function $g(P)$ is continuous on a closed region $\bar{R}$ and vanishes at every point of the boundary of R, if furthermore Gauss' theorem holds for every sphere contained in R the function $g(P)$ vanishes identically. If indeed $g(P)$ were not a constant there would be points of R at which say $g(P) > 0$ and then it would exist at least one point P_0 in R at which $g(P)$ would attain its maximum value. But this contradicts the result of the last section.

10. It follows herefrom that every function $f(P)$ for which the theorem of Gauss holds for every sphere which lies inside of a region R must be harmonic. Consider indeed the harmonic function $V(P)$ which has the same values as $f(P)$ on the surface of such a sphere. The conditions stated in § 9 apply to the function

$$g(P) = f(P) - V(P)$$

in the sphere in which it is defined. We therefore have

$$f(P) \equiv V(P)$$

which proves that $f(P)$ is harmonic.

We have so far obtained two very remarkable results: in first line we have seen that harmonic functions may be as well characterized by the property expressed by Gauss' theorem as by the differential equation $\Delta f = 0$. It is very surprising at first sight that the property of the mean involves the functions for which it holds to be analytic.

Secondly we see in using the result of § 9 that there is *at most* one function which being harmonic inside of a given region is continuous at the boundary and which has there prescribed values. Dirichlet's problem has therefore either no solution at all or just one single solution. For the sphere there exists always a solution given by Poisson's integral. In order to make a choice between these possibilities for more general regions we must study with more detail the properties of harmonic functions.

11. Harnack's theorem. We suppose in the formula (4. 4) for Poisson's integral the function $V(Q)$ not to be negative. As we have at the center O of the sphere

$$V(O) = \frac{1}{\alpha_n a^{n-1}} \int V(Q)\, do \tag{11. 1}$$

and as for any point P interior to a concentric sphere of radius $a \cdot \theta$ (the number θ being positive and less than one) we have

$$\rho \leqq a\theta, \quad (1-\theta)a \leqq r \leqq (1+\theta)a \tag{11. 2}$$

uniformly for every point Q of the surface, we can write

$$V(O)\,\frac{1-\theta^2}{(1+\theta)^n} \leqq V(P) \leqq V(O)\,\frac{1}{(1-\theta)^n}\,. \tag{11. 3}$$

Suppose now that at a point P' of the smaller sphere the inequality

$$V(P') < M \tag{11. 4}$$

holds. We then have by (11. 3)

$$V(O) < \frac{(1+\theta)^{n-1}}{(1-\theta)}\,M \tag{11. 5}$$

and consequently for every point P'' of the same sphere

$$V(P'') < \frac{(1+\theta)^{n-1}}{(1-\theta)^{n+1}}M. \tag{11. 6}$$

Furthermore we have always

$$|\, V(P'') - V(P')\,| \leqq V(0)\,\frac{(1+\theta)^{n-1} - (1-\theta)^{n+1}}{(1-\theta)(1-\theta^2)^{n-1}}\,. \tag{11. 7}$$

12. We consider now the family of functions $U(P)$ which being harmonic inside of a region are at every point of R not smaller than a given constant α and besides at a certain point P_0 of the same region not greater than a second constant β. If we apply the inequality (11. 6) to the functions $(U(P) - \alpha)$ we can easily show in using chains of spheres overlapping one another that the functions of our family have a common finite upper bound on any given closed set of points lying inside of R.

This being the case the inequality (11. 7) shows that the limiting oscillation of our family of functions vanishes at every interior point of R and that the family is therefore a normal family in the sense of *Montel*.

Consequently we can extract from every sequence U_j of functions of our family a subsequence U_{j_k} which converges continuously towards a function $V(P)$ at every point of R. The convergence must therefore be uniform on every closed subset of R and we infer herefrom that for the limiting function $V(P)$ the theorem of Gauss holds for every sphere contained in R and that $V(P)$ must therefore itself be harmonic.

We can state this result more briefly by saying that our family is not only *normal* but also *closed*; i. e. that every limiting function of functions of our family belongs itself to the same family.

The well known theorem of *Harnack* is very closely connected with these last results which are but slightly more general (though much more convenient for the applications than the older theorem).

13. Superharmonic and Subharmonic functions. Classes of functions which have been named in recent times super- and subharmonic have been used rather early in the literature. *Poincaré* has made use of them; but it was not until they were carefully studied by *F. Riesz* that their great importance for Potential theory has appeared. *Riesz* has made use of such functions under the assumption that they are semicontinuous. For our purposes it will be sufficient to use the following definition: a function $S(P)$ is called superharmonic in a region R if being continuous in that region $S(P)$ is never smaller than the mean of $S(Q)$ taken for the surfaces of every sphere of center P lying entirely in R and whose radii are less than a positive number $r_0(P)$ whose magnitude varies with P. The definition of subharmonic functions $s(P)$ is the same but for the sign of the inequality.

Calling m the lower bound of $S(P)$ in the region R one shows that if $S(P)$ be superharmonic and not a constant we have always

$$m < S(P). \tag{13. 1}$$

Taking an arbitrary point P_0 in R we construct just as in § 8 a sphere with center P_1 lying in R and possessing at its surface at least one point Q for which

(13. 2) $$S(Q) \neq S(P_1)$$

the center P_1 of the sphere having the further property that $S(P_1) = S(P_0)$. If $S(P)$ is not constant on the surface of this sphere the mean of the function taken over that sphere is actually greater than m and $S(P)$ being superharmonic we have $S(P_1) > m$. But if $S(P)$ is constant throughout the surface of the sphere we have by the condition of superharmonicity $S(P_1) \geqq S(Q)$ and by (13. 2)

(13. 3) $$S(P_1) > S(Q) \geqq m.$$

In both cases the relation (13. 1) is verified at the point P_0.

Calling M the upper bound of a subharmonic function $s(P)$ we show in like manner that either $s(P)$ is a constant or that we have for every point of R

$$s(P) < M.$$

It follows that a superharmonic function $S(P)$ which is continuous on a *closed* region $\bar{R}$ and which is not negative at the boundary of R cannot have negative values in R and that under analogous conditions a subharmonic function $s(P)$ is never positive.

14. We shall make use of some properties of the superharmonic functions to which *F. Riesz* was the first to pay attention.

Consider first a sphere lying entirely in a region R in which a superharmonic function $S(P)$ is defined and call $V(P)$ the harmonic function (calculated by Poisson's Integral) which coincides with $S(P)$ at the boundary of that sphere. Applying the reasoning at the end of the previous section to the function $(S(P) - V(P))$ which is superharmonic inside the sphere and which vanishes at its surface we have

(14. 1) $$V(P) \leqq S(P).$$

15. We consider now a function $S^*(P)$ which is equal to $S(P)$ outside of the sphere which we have just considered and equal to $V(P)$ inside of this same sphere. As we have everywhere $S^*(P) \leqq S(P)$ the characteristic property for the function $S^*(P)$ to be superharmonic is verified outside and on the boundary of our sphere. But the same holds obviously also for the points inside of the sphere as $S^*(P)$ coincides with the harmonic function $V(P)$ at these points.

Thus $S^*(P)$ is superharmonic in the whole of the region R. This property of superharmonic functions which consists in the possibility to replace such a function by another which is nowhere greater and which is harmonic in a given sphere is of paramount importance for the sequel.

16. Finally we consider two superharmonic functions $S_1(P)$ and $S_2(P)$ defined in the same region and call $S(P)$ the function which at every point of R is equal to the smaller of both numbers $S_1(P)$ and $S_2(P)$. It is readily seen that $S(P)$ is also superharmonic in R. For at a point P_0 of R we have say

$$S(P_0) = S_1(P_0)$$

and by assumption $S_1(P_0)$ is not smaller than the mean of that function taken over the surface of a sufficient small sphere with center P_0. Then $S_1(P_0)$ or which is the same $S(P_0)$ cannot be smaller than the mean taken over the surface of the same sphere of the function $S(P)$ because we have everywhere $S(P) \leqq S_1(P)$.

17. Existence of the solution of Dirichlet's Problem. We consider on the boundary $(\bar{R} - R)$ of a given region R a continuous function $\phi(Q)$ and call according to *Perron* all the functions which are continuous in $\bar{R}$ and superharmonic in R and which at every point of the boundary are not less than $\phi(Q)$ an *upper function* $S(P)$ of our problem.

In a similar way a *lower function* $s(P)$ of the problem is a function which is continuous in $\bar{R}$ never greater than $\phi(Q)$ at the boundary of R and subharmonic in that region. According to § 13 the difference $(S(P) - s(P))$ of an upper and of a lower function being superharmonic in R and not negative at the boundary is nowhere negative in R.

If we call m and M the minimum and the maximum of $\phi(Q)$ on the boundary of R the constant function equal to M is an upper function of the problem and the constant function equal to m a lower function.

Accordingly we have at every point of R and for every upper function $S(P)$

$$S(P) \geqq m \tag{17.1}$$

and the lower bound $U(P)$ of all possible upper functions at a point P of R satisfies the conditions

$$m \leqq U(P) \leqq M. \tag{17.2}$$

We will show that $U(P)$ is harmonic in R.

18. We consider a sequence of points

$$P_1, P_2, \cdots, P_j, \cdots \tag{18.1}$$

which is everywhere dense in R and for each of these points say for P_j a sequence of upper functions

(18.2) $$S_{j1}, S_{j2}, \cdots, S_{jk}, \cdots$$

which satisfy the condition

(18.3) $$U(P_j) = \lim_{k=\infty} S_{jk}(P_j).$$

We call now $S_k(P)$ the function which at every point of R is equal to the smallest of the numbers

$$M, S_{1k}(P), S_{2k}(P), \cdots, S_{kk}(P).$$

According to § 17 this function is a bounded upper function and we have

(18.4) $$U(P_j) = \lim_{k=\infty} S_k(P_j) \qquad (j = 1, 2, \cdots).$$

We replace now the functions S_k by the functions S^*_k which we have constructed in § 15 after having taken an arbitrary sphere σ inside of the region R.

The functions S^*_k are again upper functions for the problem for which relations analogous to (18.4) hold because we have at every point of R

(18.5) $$U(P) \leqq S^*_k(P) \leqq S_k(P).$$

Now by § 12 the functions $S^*_k(P)$ belong to a normal family of harmonic functions inside of the sphere σ and form a sequence which is convergent at an everywhere dense set of points in this sphere. They must therefore converge towards a harmonic function $V(P)$ and we are going to show that one must have $V(P) = U(P)$ at *all* the points of the sphere σ. This is true by construction for all those points of the countable set (18.1) which lie inside σ. We take at random another point of σ which we shall call P_0 and we add it to the points of the set (18.1). We can then construct by the same process as above a new function $V'(P)$, which being harmonic inside of σ is equal to $U'(P)$ at every point of the set (18.1) and besides at the point P_0. The functions $V(P)$ and $V'(P)$ are both continuous in σ and they are equal to one another at the points of the set (18.1) which is everywhere dense in that sphere. They must therefore coincide in σ and we have in particular $U(P_0) = V(P_0)$. This shows that $U(P)$ is harmonic in σ and as this sphere may be chosen at random the function $U(P)$ must be harmonic everywhere in R.

19. If the closed region $\bar{R}$ has the property that Dirichlet's problem is solvable for any given continuous boundary values, there exist functions $V(P; Q_0)$ of P continuous on the closed set $\bar{R}$ which vanish at a given point Q_0 of the boundary of R, which are positive at all the other boundary points of R and harmonic in this region.

This leads to the following definition: a boundary point Q_0 of a given region (for which one does not know if Dirichlet's problem is solvable or not) is called a *regular* point of the boundary if a function $V(P; Q_0)$ with the properties stated above exists.

For the case that Q_0 is a regular point of the boundary it is very easy to construct an upper function of our problem which at the point Q_0 has a value not greater than $\phi(Q_0) + \epsilon$, the positive number ϵ being chosen at random. Take a sphere $\sigma(Q_0; \epsilon)$ with center Q_0 and such that at all the points Q of the boundary of R lying inside of that sphere we have

$$\phi(Q) < \phi(Q_0) + \epsilon. \tag{19.1}$$

Call μ the necessarily positive minimum of the values which $V(P; Q_0)$ takes in $\bar{R}$ at points lying not inside $\sigma(Q_0; \epsilon)$. Then the function

$$S(P) = \phi(Q_0) + \epsilon + \frac{M - m}{\mu} V(P; Q_0)$$

is an upper function of our problem which possesses the stated property. In the same way a lower function of the problem which at the point Q_0 has a value not smaller than $\phi(Q_0) - \epsilon$ can be found.

It follows then that the harmonic function $U(P)$, which we have defined in the previous section is continuous at every regular point of the boundary of R and that it takes at such points the prescribed value $\phi(Q_0)$.

Thus for the solution of Dirichlet's problem it is not only necessary but it is also sufficient that all the points of the boundary of R be regular. It is well known that *Lebesgue* has given examples of regions—even of simply connected regions—for which not all the boundary points are regular. In the sequel we shall restrict ourselves to regions which possess a regular boundary.

Chapter II. Construction of the Solutions of Dirichlet's Problem.

20. Statement of the problem. We consider a closed bounded region $\bar{R}$ whose boundary points Q are all regular in the sense of § 19. We try to define sequences of operators

$$L_1 f, L_2 f, \cdots, L_j f, \cdots \tag{20.1}$$

which enable us to calculate by a limiting process the harmonic functions $U(P)$ whose existence we have proved in the two previous sections.

More precisely we shall take an arbitrary function $f_1(P)$ which is continuous on the closed region $\bar{R}$ and which at the boundary points Q of that region satisfies the equality

$$f_1(Q) = \phi(Q), \tag{20.1}$$

the last function being continuous on the boundary of R and taken at random. Then we shall construct a sequence of functions $f_k(P)$ given recurrently by the formulas

$$f_2(P) = L_1 f_1(P), \cdots, f_{k+1}(P) = L_k f_k(P), \cdots. \tag{20.2}$$

We shall show that the operators $L_j f$ can be chosen in such a way that the sequence $f_k(P)$ converges uniformly towards the solution $U(P)$ of Dirichlet's problem which possesses the given boundary values $\phi(Q)$.

21. We shall suppose that the operators $L_j f$ are defined for all functions $f(P)$ which are finite and continuous in the interior of R no matter if these functions can be extended to a continuous function on the closed region $\bar{R}$ or not or if these functions are bounded in R or not.

We suppose furthermore that the result $L_j f$ of each of our operations applied to functions of the above class must be a function of the same class.

Our operators $L_j f$ shall have now the following properties:

a) they are linear i. e.

$$L_j(f+g) = L_j f + L_j g$$
$$L_j(-f) = -L_j f$$

b) they are definite i. e. if $f(P) \geqq 0$ throughout R then $L_j f \geqq 0$ throughout R.

c) they do not change the value of harmonic functions i. e. if $U(P)$ is harmonic everywhere in R we have

$$L_j U(P) \equiv U(P).$$

22. If we take the operators $L_j f$ satisfying the above conditions our problem is already greatly simplified. It will indeed be sufficient to restrict the $L_j f$ in such a way that for functions $f_1(P)$ whose boundary values vanish identically the functions $f_k(P)$ considered in § 20 converge uniformly towards zero. For in the general case we can subtract from $f_1(P)$ the harmonic function $U(P)$ which has the same boundary values $\phi(Q)$ and whose existence is guaranteed and put

$$g_1(P) = f_1(P) - U(P), \quad g_{k+1}(P) = L_k g_k(P) \qquad (k = 1, 2, \cdots).$$

We have then for every value of k

$$f_k(P) = g_k(P) + U(P)$$

and if we know that the $g_k(P)$ converge towards zero the functions $f_k(P)$ must converge towards $U(P)$ as desired.

23. Superharmonic cover. Before stating the further properties of the linear operators $L_j f$ which will insure the convergence of our process we must study closer the consequences of the properties a), b) and c) of our operators.

We take an arbitrary function $f(P)$ which is continuous in the closed region $\bar{R}$ and vanishes at all the points Q of the boundary of R. We consider functions $U(P)$ harmonic everywhere in R (but not necessarily defined on the boundary of R) which satisfy the condition

$$f(P) \leqq U(P). \tag{23.1}$$

The lower bound of all these functions $U(P)$ considered at the same point P of R defines a function $S(P)$ which we shall call the *superharmonic cover* of $f(P)$.

Take a point P_0 inside of R and consider a sequence $U_j(P)$ taken among the harmonic functions defined above and such that

$$S(P_0) = \lim_{j=\infty} U_j(P_0). \tag{23.2}$$

As $f(P)$ is a bounded function it follows from (23.1) and (23.2) and from Harnack's theorem that the functions $U_j(P)$ belong to a normal family of functions. There exist therefore subsequences of the $U_j(P)$ which converge towards a function $U_{P_0}(P)$ harmonic in R and for which the following properties hold

$$S(P) \leqq U_{P_0}(P), \qquad S(P_0) = U_{P_0}(P_0). \tag{23.3}$$

From $f(P) \leqq U_{P_0}(P)$ follows that $U_{P_0}(P) \geqq 0$, if we remind ourselves of the properties of harmonic functions stated in § 9 and of the assumption that $f(P)$ vanishes at the boundary of R. From the second relation (23.3) we infer that everywhere in R we have

$$S(P) \geqq 0. \tag{23.4}$$

It is furthermore easy to verify that the superharmonic cover $S(P)$ is a continuous function. Taking in fact a sequence of points P_j which converge towards P_0 we infer on using both relations (23.3) that

$$\overline{\lim_{j=\infty}} S(P_j) \leqq S(P_0). \tag{23.5}$$

We have now

$$S(P_j) = U_{P_j}(P_j), \qquad U_{P_j}(P_0) \geqq S(P_0) \tag{23.6}$$

and on using the fact that the functions $U_{P_j}(P)$ belong to a normal family we find easily

$$\overline{\lim_{j=\infty}} S(P_j) \geqq S(P_0). \tag{23.7}$$

The continuity of $S(P)$ is thus proved by the inequalities (23.5) and (23.7).

Making now use of the assumption that any point Q_0 of the boundary of R is a regular point we can construct functions $U(P)$ which are continuous on the closed region $\bar{R}$, which are harmonic in R and not less than the given function $f(P)$ and which at the point Q_0 have a value which is smaller than a given positive number ϵ (cf. § **19**). From this and from (23.4) we infer that the superharmonic cover $S(P)$ is continuous at all the boundary points of R and that it vanishes at those points.

24. We observe now that we have

$$f(P) \leqq S(P) \leqq U_{P_0}(P) \tag{24.1}$$

and that consequently for every single operator

$$L_j f \leqq L_j S \leqq L_j U_{P_0}(P) = U_{P_0}(P).$$

It follows by the second relation (23.3) that we must have at the point P_0

$$L_j f \leqq L_j S \leqq S. \tag{24.2}$$

Now we may choose P_0 anywhere in R and see that the inequalities (24.2) hold in the whole region R.

We apply now these inequalities to the sequence of functions $f_k(P)$ defined by (20.2). Calling S_k the superharmonic cover of f_k we have

$$f_{k+1}(P) = L_k f_k(P) \leqq S_k(P)$$

and it follows

$$S_{k+1}(P) \leqq S_k(P) \tag{24.3}$$

by the definition of the superharmonic cover of $f_{k+1}(P)$.

25. The monotonic sequence of functions

$$S_1(P) \geqq S_2(P) \geqq \cdots \tag{25.1}$$

converges towards a function

$$T(P) \geqq 0. \tag{25.2}$$

We are going to show that the convergence is continuous at every point P_0 of R. Taking a sequence of points P_k converging towards P_0 we have for any $k > j$

$$S_k(P_k) \leqq S_j(P_k)$$

and therefore

$$\overline{\lim_{k=\infty}} S_k(P_k) \leqq S_j(P_0).$$

As this holds for every value of j we have

$$\overline{\lim} S_k(P_k) \leqq T(P_0). \tag{25.3}$$

We take a subsequence $S_{k_j}(P)$ of our sequence such that

$$\underline{\lim} S_k(P_k) = \lim_{j=\infty} S_{k_j}(P_{k_j}). \tag{25.4}$$

We consider functions $V_j(P)$ harmonic in R for which the following conditions hold

$$V_j(P) \geqq S_{k_j}(P), \qquad V_j(P_{k_j}) = S_{k_j}(P_{k_j}). \tag{25.5}$$

The sequence of functions $V_j(P)$ belongs to a normal family; there exists a subsequence of these functions convergent continuously in R towards a function $V(P)$ and certainly

$$V(P) \geqq T(P).$$

We have therefore finally by (25.4)

$$\underline{\lim} S_k(P_k) = V(P_0) \geqq T(P_0) \tag{25.6}$$

and the comparison with (25.3) yields

$$\lim_{k=\infty} S_k(P_k) = T(P_0).$$

It follows by the general theory that the functions $S_k(P)$ converge uniformly towards $T(P)$ on any closed subset of R. If one considers that the non-negative function $S_1(P)$ is continuous on the closed region $\bar{R}$ and vanishes on the boundary of R and if one makes use of the fact that the sequence is monotonic and that the limiting function $T(P)$ is itself non-negative it follows that the sequence (25.1) converges uniformly towards $T(P)$ on the *closed* region $\bar{R}$.

26. The limiting function $T(P)$ is furthermore restricted if among the operators $L_j f$ there are some which are *iterated*. We shall say that the operator $L_q f$ is iterated if there exists an infinite sequence of increasing integers

$$q_1 < q_2 < \cdots \tag{26.1}$$

such that the operators L_{q_j} with the indices q_i be all of them identical with L_q

$$L_q f \equiv L_{q_j} f \qquad (j = 1, 2, \cdots). \tag{26.2}$$

Given a positive number ϵ; by the result at the end of the last section there are in the sequence (26.1) numbers q_j for which

$$S_{q_j} < T + \epsilon. \tag{26.3}$$

We have now by (24.2)

$$L_{q_j} f_{q_j} \leqq L_{q_j} S_{q_j} \leqq L_{q_j}(T + \epsilon) \tag{26.4}$$

and for the last number of this relation by (26.2)

$$L_{q_j}(T + \epsilon) = L_q T + \epsilon.$$

Calling $T'_q(P)$ the superharmonic cover of $L_q T(P)$ and remembering that $S_{q_{j+1}}$ is the superharmonic cover of $L_{q_j} f_{q_j}$ we can therefore write

$$T(P) \leqq S_{q_{j+1}}(P) \leqq T'_q + \epsilon. \tag{26.5}$$

On the other hand we have for every value of the subscript k

$$L_q T \leqq L_q S_k \leqq S_k$$

and herefrom follows

$$T'_q(P) \leqq S_k(P)$$

and at the limit

$$T'_q(P) \leqq T(P). \tag{26.6}$$

Comparing with (26.5) we get finally

$$T'_q(P) \equiv T(P).$$

This result is not surprising: the operator $L_q f$ having been involved an infinite number of times already for the construction of the sequence of superharmonic functions $S_k(P)$ cannot modify any more the function $T(P)$ if we apply a construction similar to that which transforms S_q into S_{q+1}.

27. Certainly the conditions we have imposed up to now on our operators $L_j f$ are not sufficient to make the functions $f_k(P)$ converge towards zero. We could take the $L_j f$ all equal to the identical transformation and every single one of the conditions imposed so far upon them would be fulfilled.

But in that case a function $f_1(P)$ which is not identically zero would never yield a sequence $f_k(P)$ converging towards that number.

On the other hand if we can make sure that to every non-negative function $T(P) \not\equiv 0$ corresponds at least one of the iterated operators $L_q f$ for which $T'_q(P) \not\equiv T(P)$, by the result of the last section the function $T(P)$ must vanish identically and our sequence of functions $f_k(P)$ must then converge uniformly towards zero.

To prove this we consider the superharmonic covers of the functions $-f_k(P)$ and call them $-s_k(P)$. We have then

$$s_k(P) \leqq f_k(P) \leqq S_k(P) \tag{27.1}$$

and by our assumptions both of the sequences $s_k(P)$ and $S_k(P)$ must converge uniformly towards zero. The same is therefore true for the sequence of the $f_k(P)$.

28. A very general condition which we shall impose on our operators and which implies the just stated property is the following: we take a non-negative function $T(P)$ which is not a constant, which is continuous on the closed region R, which vanishes at the boundary of R and whose maximum value M is attained at an interior point P_0 of R. We assume furthermore that P_0 is such a point that there exists at least an $(n-1)$-dimensional plane α passing through P_0 and such that at every point Q of R lying on one side of α we have $T(Q) < M$. We require that under these assumptions for one at least of the iterated operators, say for $L_q f$, we have at the point P_0 itself

$$L_q T(P_0) < M. \tag{28.1}$$

We have only to prove that at a point P_0 which fulfills the above requirements we have always

$$T'_q(P_0) < M = T(P_0). \tag{28.2}$$

If there is only one point in the interior of R at which $T(P) = M$ we take this point for our P_0 and remark that by (28.1) the *convex cover* of $L_q T(P)$ lies underneath $T(P)$ at the point P_0 and that the superharmonic cover of $L_q T(P)$ lying underneath that just considered convex cover the inequality (28.2) is fulfilled.

If there be several points in R at which $T(P) = M$ these form a closed point set A lying inside of R and possessing no point on the boundary of R. We take two points P_1 and P_0 of this set whose distance from one another is

a maximum and consider the plane α passing through P_0 and normal to the vector $\overrightarrow{P_0P_1}$. All our conditions being fulfilled we must have

$$L_qT(P_0) < M \tag{28.3}$$

not only at P_0 itself but also inside a sphere σ with P_0 as center. The convex cover of $L_qT(P)$ is always less than M except perhaps on the smallest convex figure of the space which contains the points $(A - A\sigma)$. It is very easy to verify that the point P_0 is exterior to this figure and therefore we conclude just as before that (28.2) holds.

29. Our final result may be stated in the following way:

Take a bounded region R in n-space for which Dirichlet's problem has always a solution. Take a sequence of linear operators L_jf $(j = 1, 2, \cdots)$ defined for functions which are continuous in the interior of R, the result of every one of these operations being a function of the same kind. Suppose that the operators L_jf have the following properties:

1) *If $f(P) \geqq 0$ then $L_jf \geqq 0$ everywhere in R;*

2) *If $U(P)$ is harmonic everywhere in R then $L_jU \equiv U$;*

3) *A subsequence of the sequence L_jf consists of iterated operators L_qf, i. e. an infinite number of the operators L_jf are identical with L_qf;*

4) *If a non-negative function $T(P)$ which is continuous on the closed region $\bar{R}$ and which vanishes at the boundary points of R attains its maximum value $M(M > 0)$ at a point P_0 and if there is a plane α going through P_0 such that for every point Q of R lying on one side of α we have always $T(Q) < M$, then there exists at least one iterated operator L_qf such that at P_0 we have*

$$L_qT(P_0) < M. \tag{29.1}$$

Under these conditions taking an arbitrary function $f_1(P)$ continuous on the closed region $\bar{R}$ and putting

$$f_2(P) = L_1f_1(P), \cdots, f_{k+1}(P) = L_kf_k(P), \cdots \tag{29.2}$$

the sequence of functions $f_k(P)$ converges uniformly in the closed region $\bar{R}$ towards a function $U(P)$ harmonic in R and having the same boundary values as the original function $f_1(P)$.

30. Nearly all the solutions for Dirichlet's problem which have been proposed since the time of *Schwarz* and *Poincaré* and which are based on a convergent sequence of continuous operators can be treated as special cases of the foregoing theorem.

For the rest of them, as for the solution proposed by *O. D. Kellogg*, iterated operators do not appear at first sight and the conditions of the theorem of the last section must be slightly revised if it is wanted to include also these solutions in our theory. In fact suppose that among the operators $L_j f$ there be no iterated operators but that to each operator $L_j f$ we can add a finite number of operators

$$L_{j1}f, L_{j2}f, \cdots, L_{jp_j}f \tag{30.1}$$

such that we have always

$$L_{jk}(L_j f) \equiv L_j f \qquad (k = 1, 2, \cdots, p_j) \tag{30.2}$$

and that for the sequence of operators which we obtain in including all the operators (30. 1) after every $L_j f$ all the conditions of § 29 are fulfilled.

If we remark now that the sequence of functions (29. 2) has always the same limit for the original and for the completed system of operators because of the conditions (30. 2) we see that the convergence theorem holds for the $L_j f$ themselves.

31. The study of the convergence of our sequence of functions $f_k(P)$ is much more complicated if we do not choose to assume that all the points of the boundary of the region R be regular. In this connection the posthumous memoir of *O. D. Kellogg* about the converse of Gauss' Theorem will be of great help. It contains already the proof of such a convergence in a particular case. One could also think of generalizing the condition 2) of our theorem by replacing the identity by a more general relation which tends towards the former for large values of the index number j. In doing this it would perhaps be possible to include the very important and very interesting constructions of *N. Wiener* (which are discontinuous) in the general frame of our theory. But this exceeds the scope of the present note, which was to show that a fairly general result may be obtained without using any higher concept than that of normal families and that of the continuous convergence of continuous functions. It would though certainly be interesting to investigate more thoroughly, as I have done, the limits at which the theorem of convergence ceases to be true.

PASADENA, CALIFORNIA.

REFERENCES.

H. A. Schwarz, "Zur Integration der partiellen Differentialgleichung $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = 0$," *Journal für reine und angewandte Mathematik*, Bd. 74, pp. 218-253. *Ges. Abh.*, Bd. 2 (1872), pp. 175-210.

H. Poincaré, "Sur les équations aux dérivées partielles de la physique mathématique," *American Journal of Mathematics*, vol. 13 (1890).

H. Lebesgue, "Sur le problème de Dirichlet," *Comptes Rendus* (Paris), vol. 154 (1912), p. 335.

O. Perron, "Eine neue Behandlung der ersten Randwertaufgabe für $\Delta u = 0$," *Mathematische Zeitschrift*, Bd. 18 (1923), pp. 42-54.

T. Radó und F. Riesz, "Über die erste Randwertaufgabe für $\Delta u = 0$," *ibid.*, Bd. 22 (1925), pp. 41-44.

F. Riesz, "Sur les fonctions subharmoniques et leur rapport à la théorie du potentiel (Première Partie)," *Acta Mathematica*, vol. 48 (1926), pp. 329-343.

N. Wiener, "The Dirichlet problem," *Journal of Mathematics and Physics*, Massachusetts Institute of Technology, vol. 3 (1924), pp. 127-146.

O. D. Kellogg, *Foundations of Potential Theory*, Berlin, Springer, 1929.

———, "Converses of Gauss' theorem on the arithmetic mean," *Transactions of the American Mathematical Society*, vol. 36 (1934), pp. 227-242.

C. Carathéodory, "Stetige Konvergenz und normale Familien von Funktionen," *Mathematische Annalen*, Bd. 101 (1929), pp. 515-533.

BOUNDED ANALYTIC FUNCTIONS IN SEVERAL VARIABLES AND MULTIPLE LAPLACE INTEGRALS.*

By S. BOCHNER.

Introduction. We shall consider the k-dimensional Euclidean space of the real points $x = (x_1, \cdots, x_k)$ and its complex extension consisting of the complex points $z = (z_1, \cdots, z_k)$; $z_\kappa = x_\kappa + iy_\kappa$. Any point set S of the real space is the basis of a *tube* $T = T_S$ of the complex space. The latter is the point set of the complex space consisting of all k-dimensional planes

$$x_\kappa = x_\kappa^0; \qquad -\infty < y_\kappa < \infty, \qquad (\kappa = 1, \cdots, k),$$

where $(x_1^0, \cdots, x_k^0)$ is an arbitrary point of S. A tube is an (open) domain if and only if its basis is a domain. We shall say that a tube T' lies within a tube T if the closure of T' is part of the interior of T.

The transformation

$$w_\kappa = e^{-z_\kappa} \qquad (\kappa = 1, \cdots, k) \tag{1}$$

transforms a tube of the z-space into a so-called Reinhardt set of the w-space and, conversely, the image of a Reinhardt set in the w-space is a tube of the z-space.[1] For shortness, we shall term a Reinhardt region *convex* if the corresponding tube is convex in the usual sense. Any function analytic in a Reinhardt region R has a Laurent expansion

$$\sum_{-\infty}^{\infty} \cdots \sum a_{n_1 \ldots n_k} w_1^{n_1} \cdots w_k^{n_k} \tag{2}$$

which is absolutely convergent in R. On the other hand the largest domain of absolute convergence of a series of the form (2) is a convex Reinhardt region.[2] Therefore, if a function is analytic in a Reinhardt region R it also exists in the smallest convex Reinhardt region enclosing R. In z-coördinates we obtain the following result. *If a function $f(z) = f(z_1, \cdots, z_k)$ is analytic in a tube T and has the period $2\pi i$ in each variable then it also exists in the convex closure $\bar{T}$ of T and has the period $2\pi i$ in each variable.*

Similar theorems hold for many classes of analytic functions other than

---

* Received June 14, 1937.

[1] H. Behnke und P. Thullen, "Theorie der Funktionen mehrerer komplexer Veränderlichen," *Ergebnisse der Mathematik*, vol. 3 (1934), pp. 33-34.

[2] Behnke-Thullen, *ibid.*, pp. 37-38.

periodic. As an illustration we shall discuss the class of bounded functions by two different methods. In the first proof we shall represent functions of integrable square by multiple Laplace integrals of the form

$$\int_{-\infty}^{\infty} \cdots \int \exp(-z_1 t_1 - \cdots - z_k t_k) \cdot \phi(t)\, dv_t, \tag{3}$$

the symbol dv_t denoting the Euclidean volume element $dt_1 \cdots dt_k$.

Functions of integrable square. A function $f(z)$ which is analytic in T_S is of integrable square in T_S if the function

$$f_x(y) \equiv f(x_1 + iy_1, \cdots, x_k + iy_k) \tag{4}$$

belongs to the Lebesgue class L_2 over the y-space, for every $x \subset S$, and moreover,

$$\int_{-\infty}^{\infty} \cdots \int |f_x(y)|^2\, dv_y \leq K \equiv K_S. \tag{5}$$

Of course, if (5) holds only for every domain S' within S then $f(z)$ is of integrable square within T.[3]

Let $\phi(t) = \phi(t_1, \cdots, t_k)$ be a measurable function of the real t-space. If, for every $x \subset S$, the function

$$\phi_x(t) = \exp(-x_1 t_1 - \cdots - x_k t_k) \cdot \phi(t) \tag{6}$$

belongs to L_2 in the t-space then the integral (3) is absolutely and uniformly convergent within T_S and the sum function $f(z)$ is of integrable square within T_S. In proving this we may assume, applying the Heine-Borel theorem, that S is an interval

$$a_\kappa \leq x_\kappa \leq b_\kappa \qquad (\kappa = 1, \cdots, k). \tag{7}$$

Furthermore we may split the range of integration in (3) into the 2^k octants $t_1 \gtrless 0, \cdots, t_k \gtrless 0$. For instance, if we restrict ourselves to the octant $t_1 > 0, \cdots, t_k > 0$, then (3) has the value

$$\int_{-\infty}^{\infty} \cdots \int \exp[-(y_1 t_1 + \cdots + y_k t_k) i] \cdot \phi_x(t)\, dv_t. \tag{8}$$

---

[3] A function has a property within a tube T if it has the property in every tube T' within T.

The absolute and uniform convergence of (8) follows immediately from the relation

$$\left[\int_0^\infty \cdots \int |\phi_{x+\xi}(t)|\, dv_t\right]^2$$

$$\leq \int_0^\infty \cdots \int \exp(-2\xi_1 t_1 - \cdots - 2\xi_k t_k)\, dv_t \cdot \int_0^\infty \cdots \int |\phi_x(t)|^2\, dv_t$$

which is a consequence of Schwarz' inequality. The sum function of (8) is therefore analytic in $z_1, \cdots, z_k$, and relation (5) follows immediately from the Parseval equation

$$\int_{-\infty}^\infty \cdots \int |f_x(y)|^2\, dv_y = (2\pi)^{k/2} \int_0^\infty \cdots \int |\phi_x(t)|^2\, dv_t.$$

Any other octant can be treated in a similar fashion or can be reduced to the first octant by changing the signs of the negative t-components.

Conversely, any function which is of integrable square in S has a unique representation of the form (3). In fact, by Plancherel's theorem, the function $f_x(y)$ has a representation

$$(9) \qquad f_x(y) \sim \int_{-\infty}^\infty \cdots \int \exp[-(y_1 t_1 + \cdots + y_k t_k)i] \cdot \phi_x(t)\, dv_t,$$

and we have to prove that (6) holds. Integrating (5) with respect to x over the domain S we obtain

$$\int_T |f_x(y)|^2\, dv_y dv_x < \infty.$$

Hence there exists a constant A such that for any polycylinder

$$P_\zeta : |z_\kappa - \zeta_\kappa| \leq r_\kappa$$

which lies entirely in T_S the relation

$$\int_{P\zeta} |f(z)|^2\, dv_x dv_y \leq A$$

holds. Expanding $f(z)$ in a power series around ζ we easily obtain

$$|f(\zeta)|^2 \leq \frac{1}{\pi^k r_1^2 \cdots r_k^2} \int_{P\zeta} |f(z)|^2\, dv_x dv_y.$$

Thus $f(z)$ is bounded within S. With the Fejer kernel

$$F_n(\eta) = \prod_{\kappa=1}^{k} \frac{(\sin (n/2)\eta_\kappa)^2}{\pi n \eta_\kappa^2}$$

we form the sequence of functions

$$(10) \quad f^n(z) = \int_{-\infty}^{\infty} \cdots \int f(x_1 + iy_1 + i\eta_1 + , \cdots, x_k + iy_k + i\eta_k) F_n(\eta_1, \cdots, \eta_k) dv_\eta.$$

They have the following properties: α) they are analytic and of integrable square within T_S, β) for each (x, y)

$$\lim_{n\to\infty} f_x^n(y) = f_x(y),$$

γ) the Fourier transform $\phi_x^n(t)$ of $f_x^n(t)$, see (9), has the value

$$\phi_x(t) \prod_{\iota=1}^{k} \left(1 - \frac{|t_\kappa|}{n}\right),$$

if $|t_\kappa| \leq 1$ and the value 0 if some t_κ is > 1.[4] As a consequence of these properties it is sufficient to prove relation (6) for functions $\phi_x(t)$ vanishing outside a finite interval of the t-space. We pick any special point $x^0 \subset S$ and introduce the functions

$$(11) \quad \phi(t) = \exp(x_1^0 t_1 + \cdots + x_k^0 t_k) \cdot \phi_{x^0}(t)$$

$$g(z) = \int_{-\infty}^{\infty} \cdots \int \exp(-z_1 t_1 - \cdots - z_k t_k) \cdot \phi(t) dv_t.$$

Since $\phi(t)$ vanishes outside a finite interval the function $g(z)$ is analytic. By (11), it is equal to the given function $f(z)$ on the linear manifold $x_1 = x_1^0, \cdots, x_k = x_k^0$. Therefore $f(z) = g(z)$ identically, and this completes the proof of (6).

Finally, the point set S for which $\phi_x(t)$ belongs to L_2 in the t-space is convex. In fact, if x^1 and x^2 are two points of S, and $x = \alpha x^1 + (1 - \alpha)x^2$, $0 < \alpha < 1$, is a point of the segment joining these points, then the quantities

$$E_1 = \exp(-x_1^1 t_1 - \cdots - x_k^1 t_k), \qquad E_2 = \exp(-x_1^2 t_1 - \cdots - x_k^2 t_k)$$
$$E = \exp(-x_1 t_1 - \cdots - x_k t_k)$$

[4] S. Bochner, *Vorlesungen über Fouriersche Integrale*, 1932, p. 191, Satz 57.

satisfy the relation

$$E = E_1^{\alpha} E_2^{1-\alpha} \leq \max(E_1, E_2) \leq E_1 + E_2.$$

Hence we conclude that *a function which is analytic and of integrable square within a tube exists and is of integrable square within the convex closure of the tube.*

Bounded functions. From the last statement we shall conclude that *a function which is analytic and bounded within a tube exists and is bounded within its convex closure.*

For the proof we may assume that S is a bounded domain whose closure is contained in the octant $x_1 > 0, \cdots, x_k > 0$. This being the case, the function

$$g(z) = \frac{f(z)}{z_1 \cdots z_k} \tag{12}$$

is of integrable square within T and therefore it exists and is bounded within $\tilde{T}$. Thus $f(z)$ is bounded within T and $O(z_1 \cdots z_k)$ within T. Given any two points x^1 and x^2 of S and any real numbers $\eta_1, \cdots, \eta_k$ we form the function of the one complex variable $w = u + iv$,

$$h(w) = f(x_1^1 + w(x_1^2 - x_1^1) + i\eta_1, \cdots, x_k + w(x_k^2 - x_k^1) + i\eta_k).$$

It is defined in the strip $0 \leq u \leq 1$, it is bounded on the boundary $u = 0$, $u = 1$, and it is $O(v^k)$ as $v \to \pm \infty$. By Phragmen-Lindelöf, $h(w)$ is bounded in the strip, and any upper bound for its values on the boundary is a bound for its values in the interior of the strip. For appropriate values of x^1, x^2, η_k and w every point of $\tilde{T}$ can be reached, and hence we easily conclude that $f(z)$ is bounded within $\tilde{T}$.

Functions in octant-shaped tubes. If a Reinhardt region contains the origin and is star-shaped with respect to the origin then the corresponding tube is "octant-shaped": if $(x_1^0, \cdots, x_k^0)$ is a point of its base S then the whole octant

$$x_\kappa \leq x_\kappa^0 \qquad (\kappa = 1, \cdots, k) \tag{13}$$

is contained in S. The convex closure of an octant-shaped tube is again octant-shaped.

If $f(z)$ is analytic in an octant-shaped tube T it also exists in the convex closure $\tilde{T}$ of T; if it is bounded within T it is also bounded within $\tilde{T}$.

Let $x' = (x'_1, \cdots, x'_k)$, $x'' = (x'_1, \cdots, x''_k)$ be any two points of S. We want to show that $f(z)$ is analytic for the points

$$x_\kappa \leq \alpha x'_\kappa + (1-\alpha)x''_\kappa, \qquad 0 \leq \alpha \leq 1 \tag{14}$$

of the basis S of T.

For $\epsilon > 0$ sufficiently small and $A > 0$ sufficiently large the two polycylinders

$$\begin{aligned} P' &: \quad |z_\kappa + A| \leq x'_\kappa + 2\epsilon + A \\ P'' &: \quad |z_\kappa + A| \leq x''_\kappa + 2\epsilon + A \end{aligned} \tag{15}$$

are both contained in T. The smallest convex Reinhardt region (*not* tube) of the z-space which is concentric with, and contains P' and P'', contains the polycylinders

$$|z_\kappa + A| \leq (x'_\kappa + \epsilon + A)^a (x''_\kappa + \epsilon + A)^{1-a}. \tag{16}$$

The sum of these point-sets for variable values of A contains the points of S for which

$$x_\kappa < \alpha(x'_\kappa + \epsilon) + (1-\alpha)(x''_\kappa + \epsilon)$$

and therefore $f(z)$ is analytic in (14). The complex variables $z_1, \cdots, z_k$ can be altered by arbitrary fixed quantities $iy_1^0, \cdots, iy_k^0$; therefore $f(z)$ is analytic for all points whose real parts satisfy relation (14). The points x', x'' being arbitrary points of S we thus conclude that $f(z)$ is analytic throughout T. This proves the first part of the theorem. The second part is a consequence of the following remark.[5] If $|f(z)|$ is $\leq M$ in (15) then it is $\leq M \cdot C$ in (16) where C depends only on x'_κ, x''_κ, ϵ, A.

Application to bounded functions. We shall give a new proof of our previous statement concerning bounded functions, and we shall again assume that our function is a function of integrable square, see (12).

If a function $f(z)$ of one variable is of integrable square in the closed interval

$$a \leq x \leq b$$

then Cauchy's formula

$$f(z) = \frac{1}{2\pi i}\int_{a+i\infty}^{a-i\infty} \frac{f(s)}{s-z}\,dx + \frac{1}{2\pi i}\int_{b-i\infty}^{b+i\infty} \frac{f(s)}{s-z}\,ds$$

[5] The remark can be proved easily by the reasoning in H. Tietze, "Über den Bereich absoluter Konvergenz von Potenzreihen mehrerer Veränderlichen," *Mathematische Annalen* 99 (1928), p. 181.

leads to a decomposition of $f(z)$ into a sum of two functions $f_1(z)$, $f_2(z)$ of which $f_1(z)$ is analytic and bounded within $x > a$ and $f_2(z)$ is analytic and bounded within $x < b$. Up to additive constants such a decomposition is unique.[6] More generally, if $f(z) = f(z_1, \cdots, z_k)$ is of integrable square in the tube (7), it can be written as a sum of 2^k functions $f_\nu(z)$, $(\nu = 1, \cdots, 2^k)$, each of which is analytic and bounded within an octant containing the interior of the tube (7). For example, if $k = 2$, the four octants are 1) $x_1 > a_1$, $x_2 > a_2$, 2) $x_1 > a_1$, $x_2 < b_2$, 3) $x_1 < b_1$, $x_2 > a_2$, 4) $x_1 < b_1$, $x_2 < b$. The functions $f_\nu(z)$ are again uniquely determined up to additive constants.

We cover the basis S by a denumerable number of octants

$$a_\kappa{}^n \leq x_\kappa \leq b_\kappa{}^n, \qquad (n = 1, 2, \cdots).$$

By the uniqueness property the functions $f_\nu{}^n(z)$ can be so normalized that for fixed ν they are analytic continuations of each other. Therefore $f(z)$ is a sum of functions $f_\nu(z)$ each of which is analytic and bounded within an octant-shaped tube T_ν containing T. By the result of the previous section each $f_\nu(z)$ is analytic, and bounded within the convex closure $\tilde{T}_\nu$ of T_ν. But the intersection of the domains $\tilde{T}_\nu$ contains the convex closure $\tilde{T}$ of T.

PRINCETON UNIVERSITY.

[6] Harald Bohr, "Fastperiodische Funktionen III," *Acta Mathem.*, Bd. 47 (1926), pp. 250-251.

THE CHARACTERS OF THE SYMMETRIC GROUP.*

By F. D. MURNAGHAN.

We have given in a recent paper (**1**) a recurrence formula (generalising the classical formula of Schur) which makes it possible to determine the characters of the symmetric group on n letters when those of the symmetric groups on a lesser number m letters ($m = 1, 2, \cdots, n-1$) are known. This method, whilst quite convenient for a systematic construction of the character tables of the symmetric groups, labors under the handicap of all recurrence methods: if one wishes to know the characters of a particular symmetric group (say $n = 20$) one must first know the characters of the symmetric groups on $m = 1, 2, \cdots 19$ letters. Explicit and convenient formulae for the characters of *certain* classes (e. g. the identity class and the transposition class) are available but the number of these classes is very restricted; thus for the class $\alpha_1 = n-4$, $\alpha_2 = 2$ (i. e. the class consisting of two transpositions, which is important in the applications to nuclear physics) it is not easy to give a convenient explicit formula. The best one can do at present is to use the recurrence formula mentioned above together with the explicit formula of Frobenius for the characters of the transposition class ($\alpha_1 = m-2, \alpha'_2 = 1$) of the symmetric group on $m = n-2$ letters. It is, however, possible to attack the problem of furnishing explicit formulae for the characters of the symmetric group from another side: instead of furnishing formulae for the characters of a given *class* (the *representation* varying) we may give formulae for the characters of a given *representation* (the *class* varying). This point of view has certain advantages when we are concerned with the applications to nuclear physics; for, there, only certain representations (namely, those corresponding to partitions of n into not more than four parts) play a rôle. Certain formulae of this type were given by Frobenius (**2**) and these were extended by Littlewood and Richardson (**3**). The method followed by the last named writers is somewhat complicated and it is the purpose of the present paper to show that the formulae they give may be readily obtained by the methods of our previous paper and to further extend their results. We shall suppose the reader to be acquainted with the notation and methods of the paper (**1**) and shall refer to this paper, where necessary, by merely giving the page number.

---

* Received August 30, 1937.

1. The characters of the two-element-partition representations $\mathbf{D}(\lambda_1, \lambda_2)$. If $(\lambda) = (\lambda_1, \lambda_2, \cdots, \lambda_k)$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ is any partition of n the characteristic of the reducible representation $\Delta(\lambda)$ of the symmetric group on n letters is the product $q_{\lambda_1}(\boldsymbol{s}) \cdots q_{\lambda_k}(\boldsymbol{s})$ where the $q_{\lambda_j}(\boldsymbol{s})$, $j = 1, 2, \cdots, k$, are the various principal characteristics (pp. 444-445). From the formula for $q_n(\boldsymbol{s})$: $q_n(\boldsymbol{s}) = \sum_{(\alpha)} \frac{1}{\alpha_1 ! \, \alpha_2 ! \cdots \alpha_n !} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$ it is clear that the coefficient of $\frac{1}{\alpha_1 ! \cdots \alpha_n !} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$ in the product $s_1 \frac{\partial q_n}{\partial s_1}$ (or, equivalently, $s_1 q_{n-1}$, p. 445) is α_1. Equally evident is the fact that the coefficient of the same expression in the product $s_1{}^2 \frac{\partial^2 q_n}{\partial s_1{}^2}$ (or, equivalently, $s_1{}^2 q_{n-2}$) is $\alpha_1(\alpha_1 - 1)$; and, generally, the coefficient of the expression referred to in $s_1{}^{\beta_1} q_{n-\beta_1}$ is $\alpha_1(\alpha_1 - 1) \cdots (\alpha_1 - \beta_1 + 1)$. Similarly the coefficient of the same expression in $\left(\frac{s_2}{2}\right)^{\beta_2} q_{n-2\beta_2}$ is $\alpha_2(\alpha_2 - 1) \cdots (\alpha_2 - \beta_2 + 1)$; and, generally, the coefficient of $\frac{1}{\alpha_1 ! \cdots \alpha_n !} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$ in $\left(\frac{s_j}{j}\right)^{\beta_j} q_{n-j\beta_j}$ is $\alpha_j(\alpha_j - 1) \cdots (\alpha_j - \beta_j + 1)$. Hence the coefficient of this same expression in $\left(\frac{s_1}{1}\right)^{\beta_1} \cdots \left(\frac{s_n}{n}\right)^{\beta_n} q_{n-\lambda_2}$, where $\lambda_2 = \beta_1 + 2\beta_2 + \cdots + n\beta_n$, is the product

$$\prod_{j=1}^{n} \{\alpha_j(\alpha_j - 1) \cdots (\alpha_j - \beta_j + 1)\}.$$

Since

$$q_{\lambda_2}(\boldsymbol{s}) = \sum_{(\beta)} \frac{1}{\beta_1 ! \cdots \beta_n !} \left(\frac{s_1}{1}\right)^{\beta_1} \cdots \left(\frac{s_n}{n}\right)^{\beta_n}$$

it follows that the coefficient of

$$\frac{1}{\alpha_1 ! \, \alpha_2 ! \cdots \alpha_n !} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$$

in the product $q_{\lambda_1}(s) q_{\lambda_2}(\boldsymbol{s})$, where $\lambda_1 + \lambda_2 = n$, is the summation over all classes (β) of the symmetric group on λ_2 letters of the product $\binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}$ $\left(\text{where } \binom{a}{b} \text{ is the binomial coefficient } \frac{a!}{a-b\,!\,b!}\right)$. In other words the character of the class (α) in the reducible representation $\Delta(\lambda_1, \lambda_2)$ is the sum $\sum_{(\beta)} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}$. Since the characteristic of the irreducible representation $D(\lambda_1, \lambda_2)$ is the two rowed determinant $\begin{vmatrix} q_{\lambda_1}(\boldsymbol{s}) q_{\lambda_1+1}(\boldsymbol{s}) \\ q_{\lambda_2-1}(\boldsymbol{s}) q_{\lambda_2}(\boldsymbol{s}) \end{vmatrix}$ (p. 460) we have

$D(\lambda_1, \lambda_2) = \Delta(\lambda_1, \lambda_2) - \Delta(\lambda_1 + 1, \lambda_2 - 1)$ and so the character of the class (α) in the irreducible representation $D(\lambda_1, \lambda_2)$ is

$$\sum_{(\beta)} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} - \sum_{(\gamma)} \binom{\alpha_1}{\gamma_1} \cdots \binom{\alpha_n}{\gamma_n}$$

the summations being over all classes (β) and (γ) of the symmetric groups on λ_2 and $\lambda_2 - 1$ letters, respectively. Corresponding to each class (γ) of the symmetric group on $\lambda_2 - 1$ letters is the class $\beta_1 = \gamma_1 + 1$, $\beta_2 = \gamma_2, \cdots$, $\beta_n = \gamma_n$ of the symmetric group on λ_2 letters and so we may write our formula for the character of the class (α) in the irreducible representation $D(\lambda_1, \lambda_2)$ as follows:

$$(1) \qquad D(\lambda_1, \lambda_2)_{(\alpha)} = \sum_{(\gamma)} \frac{1}{\gamma_1 + 1!}\, \alpha_1(\alpha_1 - 1) \cdots$$
$$\cdots (\alpha_1 - \gamma_1 + 1)(\alpha_1 - 2\gamma_1 - 1) \binom{\alpha_2}{\gamma_2} \cdots \binom{\alpha_n}{\gamma_n} + \sum_{(\beta)} \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n}$$

the summation in the first term being over all classes (γ) of the symmetric group on $\lambda_2 - 1$ letters and in the second over all *those classes* (β) *of the symmetric group on* λ_2 *letters which do not contain any unary cycle* (= *fixed letter*).

As examples of the use of this formula we list the following:

$D(n-1, 1)_{(\alpha)} = \alpha_1 - 1$ (this degenerate case being best read off directly from the relation $\{n-1, 1\} = q_{n-1}q_1 - q_n$)

$D(n-2, 2)_{(\alpha)} = \frac{1}{2}\alpha_1(\alpha_1 - 3) + \alpha_2$

$D(n-3, 3)_{(\alpha)} = \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5) + (\alpha_1 - 1)\alpha_2 + \alpha_3$

$D(n-4, 4)_{(\alpha)} = \frac{1}{24}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 7) + \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_2$
$+ (\alpha_1 - 1)\alpha_3 + \frac{1}{2}\alpha_2(\alpha_2 - 1) + \alpha_4$

$D(n-5, 5)_{(\alpha)} = \frac{1}{120}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 9)$
$+ \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_2 + \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_3$
$+ \frac{1}{2}(\alpha_1 - 1)\alpha_2(\alpha_2 - 1) + (\alpha_1 - 1)\alpha_4 + \alpha_2\alpha_3 + \alpha_5$

$D(n-6, 6)_{(\alpha)} = \frac{1}{720}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 11)$
$+ \frac{1}{24}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 7)\alpha_2$
$+ \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_3 + \frac{1}{4}\alpha_1(\alpha_1 - 3)\alpha_2(\alpha_2 - 1)$
$+ \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_4 + (\alpha_1 - 1)\alpha_2\alpha_3 + (\alpha_1 - 1)\alpha_5$
$+ \frac{1}{6}\alpha_2(\alpha_2 - 1)(\alpha_2 - 2) + \alpha_2\alpha_4 + \frac{1}{2}\alpha_3(\alpha_3 - 1) + \alpha_6$

$D(n-7, 7)_{(\alpha)} = (1/7!)\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 5)(\alpha_1 - 13)$
$+ \frac{1}{120}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 9)\alpha_2$
$+ \frac{1}{24}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 7)\alpha_3$
$+ \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_4 +$ (see next page)

$$
\begin{aligned}
&+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2(\alpha_2-1) + \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_3 \\
&+ \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_5 + (\alpha_1-1)\alpha_6 + (\alpha_1-1)\alpha_2\alpha_4 \\
&+ \tfrac{1}{6}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2) \\
&+ \tfrac{1}{2}(\alpha_1-1)\alpha_3(\alpha_3-1) + \alpha_2\alpha_5 + \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3 \\
&+ \alpha_3\alpha_4 + \alpha_7.
\end{aligned}
$$

$$
\begin{aligned}
D(n-8,8)_{(\alpha)} = {} & (1/8!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-5)(\alpha_1-6)(\alpha_1- \\
&+ (1/6!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-11)\alpha_2 \\
&+ (1/5!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-9)\alpha_3 \\
&+ \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_4 \\
&+ \tfrac{1}{48}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_2(\alpha_2-1) \\
&+ \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2\alpha_3 + \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_5 \\
&+ \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_6 + \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_4 \\
&+ \tfrac{1}{12}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)(\alpha_2-2) + \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_3(\alpha_3-1) \\
&+ (\alpha_1-1)\alpha_2\alpha_5 + \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3 + (\alpha_1-1)\alpha_3\alpha_4 \\
&+ (\alpha_1-1)\alpha_7 + \tfrac{1}{24}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-3) \\
&+ \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1) + \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_4 + \alpha_2\alpha_6 + \alpha_3\alpha_5 \\
&+ \tfrac{1}{2}\alpha_4(\alpha_4-1) + \alpha_8.
\end{aligned}
$$

It is worthy of note that these *formulae* are the same for all values of n; their *evaluations*, of course, depend on n since (α) is a class of the symmetric group on n letters, i. e. $\alpha_1 + 2\alpha_2 + \cdots + n\alpha_n = n$. The formula for $D(\lambda_1, \lambda_2)_{(\alpha)}$ involves only the $\alpha_1, \cdots, \alpha_{\lambda_2}$ being independent of the α_j for which $j > \lambda_2$. The first term in each formula yields, when α_1 is put equal to n, Frobenius' formula for the dimension of the representation $D(\lambda_1, \lambda_2)$. As a check on the accuracy of the formulae (e. g., to detect a possible error in copying) we may observe that $D(\lambda_1, \lambda_2) = 0$ when $\lambda_1 = \lambda_2 - 1$, i. e. when $n = 2\lambda_2 - 1$; hence the expression for $D(\lambda_1, \lambda_2)_{(\alpha)}$ must vanish when α is *any* class of $2\lambda_2 - 1$. E. g., the polynomial in $(\alpha_1, \alpha_2, \alpha_3)$ giving $D(n-3, 3)_{(\alpha)}$ must vanish when $(\alpha_1, \alpha_2, \alpha_3)$ have the following sets of values: $(5,0,0)$, $(3,1,0)$, $(2,0,1)$, $(1,2,0)$, $(0,1,1)$. Since the characters of any representation follow from those of the associated representation by a mere change of sign of the characters of the odd classes the formulae derived above furnish the characters of all representations which correspond to partitions of n in which no element > 2.

2. The characters of the three-element-partition representations $D(\lambda_1, \lambda_2, \lambda_3)$. On expanding the characteristic

$$
\{\lambda_1, \lambda_2, \lambda_3\} = \begin{vmatrix} q_{\lambda_1}(s) & q_{\lambda_1+1}(s) & q_{\lambda_1+2}(s) \\ q_{\lambda_2-1}(s) & q_{\lambda_2}(s) & q_{\lambda_2+1}(s) \\ q_{\lambda_3-2}(s) & q_{\lambda_3-1}(s) & q_{\lambda_3}(s) \end{vmatrix}
$$

of $D(\lambda_1, \lambda_2, \lambda_3)$ in terms of the first row and observing that

$$\begin{vmatrix} q_{\lambda_2-1}(s) & q_{\lambda_2+1}(s) \\ q_{\lambda_3-2}(s) & q_{\lambda_3}(s) \end{vmatrix} = \begin{vmatrix} q_{\lambda_2-1}(s) & q_{\lambda_2}(s) \\ q_{\lambda_3-1}(s) & q_{\lambda_3}(s) \end{vmatrix} + \begin{vmatrix} q_{\lambda_2}(s) & q_{\lambda_2+1}(s) \\ q_{\lambda_3-2}(s) & q_{\lambda_3-1}(s) \end{vmatrix}$$

$$= \{\lambda_2 - 1, \lambda_3\} + \{\lambda_2, \lambda_3 - 1\} = \frac{\partial}{\partial s_1}\{\lambda_2, \lambda_3\}$$

we find

$$\{\lambda_1, \lambda_2, \lambda_3\} = q_{\lambda_1}(s)\{\lambda_2, \lambda_3\} - q_{\lambda_1+1}(s)\frac{\partial}{\partial s_1}\{\lambda_2, \lambda_3\} + q_{\lambda_1+2}(s)\{\lambda_2 - 1, \lambda_3 - 1\}.$$

If, now, the characters of the symmetric groups on $\lambda_2 + \lambda_3$ and $\lambda_2 + \lambda_3 - 2$ letters are known we write

$$\{\lambda_2, \lambda_3\} = \sum_{(\beta)} \frac{\{\lambda_2, \lambda_3\}_{(\beta)}}{\beta_1! \cdots \beta_n!} \left(\frac{s_1}{1}\right)^{\beta_1} \cdots \left(\frac{s_n}{n}\right)^{\beta_n}$$

the summation being over all classes (β) of the symmetric group on $\lambda_2 + \lambda_3$ letters. Hence the coefficient of

$$\frac{1}{\alpha_1! \cdots \alpha_n!} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$$

in $q_{\lambda_1}(s)\{\lambda_2, \lambda_3\}$ is

$$\sum_{(\beta)} \{\lambda_2, \lambda_3\}_{(\beta)} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n}$$

Also

$$\frac{\partial}{\partial s_1}\{\lambda_2, \lambda_3\} = \sum_{(\beta')} \frac{\{\lambda_2, \lambda_3\}_{(\beta)}}{\beta'_1!\,\beta_2! \cdots \beta_n!} \left(\frac{s_1}{1}\right)^{\beta'_1} \left(\frac{s_2}{2}\right)^{\beta_2} \cdots \left(\frac{s_n}{n}\right)^{\beta_n}$$

where $\beta'_1 = \beta_1 - 1$ so that $(\beta') = (\beta'_1, \beta_2, \cdots, \beta_n)$ is an arbitrary class of the symmetric group on $\lambda_2 + \lambda_3 - 1$ letters; hence the coefficient of

$$\frac{1}{\alpha_1! \cdots \alpha_n!} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$$

in the product $q_{\lambda_1+1}(s)\frac{\partial}{\partial s_1}\{\lambda_2, \lambda_3\}$ is

$$\sum_{(\beta')} \{\lambda_2, \lambda_3\}_{(\beta)} \binom{\alpha_1}{\beta'_1} \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n}$$

the summation being over all classes (β') of the symmetric group on $\lambda_2 + \lambda_3 - 1$ letters and the coefficients $\{\lambda_2, \lambda_3\}_{(\beta)}$ being the characters of $D(\lambda_2, \lambda_3)$ corresponding to those classes of the symmetric group on $\lambda_2 + \lambda_3$ letters which contain *at least one* fixed letter. Finally the coefficient of

$$\frac{1}{\alpha_1! \cdots \alpha_n!} \left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$$

in the product $q_{\lambda_1+2}(s)\{\lambda_2-1, \lambda_3-1\}$ is

$$\sum_{(\gamma)} \{\lambda_2-1, \lambda_3-1\}_{(\gamma)} \binom{\alpha_1}{\gamma_1} \cdots \binom{\alpha_n}{\gamma_n}$$

the summation being over all classes (γ) of the symmetric group on $\lambda_2+\lambda_3-2$ letters. Hence we have the result

$$\{\lambda_1, \lambda_2, \lambda_3\}_{(\alpha)} = \sum_{(\beta)} \{\lambda_2, \lambda_3\}_{(\beta)} \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_n}{\beta_n} - \sum_{(\beta')} \{\lambda_2, \lambda_3\}_{(\beta)} \binom{\alpha_1}{\beta'_1} \cdots \binom{\alpha_n}{\beta_n} + \sum_{(\gamma)} \{\lambda_2-1, \lambda_3-1\}_{(\gamma)} \binom{\alpha_1}{\gamma_1} \cdots \binom{\alpha_n}{\gamma_n}.$$

The difference of the first two terms may be written as

$$\sum_{(\beta')} \frac{1}{\beta_1!} \{\lambda_2, \lambda_3\}_{(\beta)} \alpha_1(\alpha_1-1) \cdots (\alpha_1-\beta'_1+1)(\alpha_1-2\beta'_1-1) \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n} + \sum_{(\delta)} \{\lambda_2, \lambda_3\}_{(\delta)} \binom{\alpha_2}{\delta_2} \cdots \binom{\alpha_n}{\delta_n}$$

where the summation with respect to (δ) is over all classes (δ) of the symmetric group on $\lambda_2+\lambda_3$ letters *which do not contain any unary cycle* (= fixed letter). Hence we have the three-element-partition formula:

(2) $D(\lambda_1, \lambda_2, \lambda_3)_{(\alpha)}$

$$= \sum_{(\beta')} \frac{1}{\beta_1!} \{\lambda_2, \lambda_3\}_{(\beta)} \alpha_1(\alpha_1-1) \cdots (\alpha_1-\beta'_1+1)(\alpha_1-2\beta'_1-1) \binom{\alpha_2}{\beta_2} \cdots + \sum_{(\delta)} \{\lambda_2, \lambda_3\}_{(\delta)} \binom{\alpha_2}{\delta_2} \cdots \binom{\alpha_n}{\delta_n} + \sum_{(\gamma)} \{\lambda_2-1, \lambda_3-1\}_{(\gamma)} \binom{\alpha_1}{\gamma_1} \cdots \binom{\alpha_n}{\gamma_n}.$$

As an illustration of the application of this formula we consider the simplest case: $D(n-2, 1^2)$. Here $\lambda_2=1$, $\lambda_3=1$ so that $\lambda_2+\lambda_3-1=1$ and there is only one class (β'), namely, $\beta'_1=1$. There is only one class (δ), namely, $\delta_2=1$ and the last term yields unity. Hence

$$D(n-2, 1^2)_{(\alpha)} = \tfrac{1}{2}\alpha_1(\alpha_1-3) - \alpha_2 + 1 = \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2) - \alpha_2.$$

As examples of the use of (2) we list the following:

$$\begin{aligned}
D(n-2, 1^2)_{(\alpha)} &= \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2) - \alpha_2 \\
D(n-3, 2, 1)_{(\alpha)} &= \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4) - \alpha_3 \\
D(n-4, 3, 1)_{(\alpha)} &= \tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6) \\
&\quad + \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2 - \tfrac{1}{2}\alpha_2(\alpha_2-3) - \alpha_4 \\
D(n-4, 2^2)_{(\alpha)} &= \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5) - (\alpha_1-1)\alpha_3 \\
&\quad + \alpha_2(\alpha_2-2)
\end{aligned}$$

$$D(n-5,4,1)_{(\alpha)} = \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-8)$$
$$+ \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_2$$
$$+ \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_3 - (\alpha_2-1)\alpha_3 - \alpha_5$$

$$D(n-5,3,2)_{(\alpha)} = \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-5)(\alpha_1-7)$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2 - \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_3$$
$$- (\alpha_1-1)\alpha_4 + \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1) + \alpha_2\alpha_3$$

$$D(n-6,5,1)_{(\alpha)} = \tfrac{1}{144}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)(\alpha_1-10)$$
$$+ \tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_2$$
$$+ \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_3 + \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_4$$
$$+ \tfrac{1}{4}(\alpha_1-1)(\alpha_1-2)\alpha_2(\alpha_2-1)$$
$$- \tfrac{1}{6}\alpha_2(\alpha_2-1)(\alpha_2-2) - \alpha_2\alpha_4 - \tfrac{1}{2}\alpha_3(\alpha_3-1) - \alpha_6$$

$$D(n-6,4,2)_{(\alpha)} = \tfrac{1}{80}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-6)(\alpha_1-9)$$
$$+ \tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_2$$
$$- \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_4 + \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)$$
$$-(\alpha_1-1)\alpha_5 + \tfrac{1}{2}\alpha_2(\alpha_2-1)(\alpha_2-3) + \alpha_2\alpha_4$$

$$D(n-6,3^2)_{(\alpha)} = \tfrac{1}{144}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-7)(\alpha_1-8)$$
$$+ \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_2$$
$$- \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_3 - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_4$$
$$+ \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1) + \alpha_1(\alpha_2-1)\alpha_3$$
$$- \tfrac{1}{2}\alpha_2(\alpha_2-1)(\alpha_2-4) - \alpha_2\alpha_3 - \alpha_2\alpha_4 + \alpha_3(\alpha_3-1)$$

$$D(n-7,6,1)_{(\alpha)} = \tfrac{1}{840}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-6)(\alpha_1-12)$$
$$+ \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-8)\alpha_2$$
$$+ \tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_3$$
$$+ \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_4$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_2(\alpha_2-1)$$
$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_2\alpha_3$$
$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_5$$
$$- \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3 - \alpha_2\alpha_5 - \alpha_3\alpha_4 - \alpha_7$$

$$D(n-7,5,2)_{(\alpha)} = \tfrac{1}{360}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)(\alpha_1-11)$$
$$+ \tfrac{1}{20}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)\alpha_2$$
$$+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_3$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2(\alpha_2-2) - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_5$$
$$+ \tfrac{1}{3}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$- \tfrac{1}{2}(\alpha_1-1)\alpha_3(\alpha_3-1) - (\alpha_1-1)\alpha_6$$
$$+ \alpha_2(\alpha_2-2)\alpha_3 + (\alpha_2-1)\alpha_5$$

$$D(n-7,4,3)_{(\alpha)} = \tfrac{1}{360}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-8)(\alpha_1-10)$$
$$+ \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-8)\alpha_2$$
$$- \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_3$$
$$- \tfrac{1}{2}\alpha_1(\alpha_1-1)\alpha_3 - \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_4$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_2(\alpha_2-1) + \text{(see next page)}$$

$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_2\alpha_3 - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_5$$
$$+ (\alpha_1-1)(\alpha_3-1)\alpha_3 - \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3 - \alpha_2\alpha_5 + \alpha_3\alpha_4.$$

$$D(n-8,7,1)_{(a)} = (1/8.6!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-5)(\alpha_1-7)(\alpha_1$$
$$+ \tfrac{1}{144}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)(\alpha_1-10)\alpha_2$$
$$+ \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-8)\alpha_3$$
$$+ \tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_4$$
$$+ \tfrac{1}{16}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_2(\alpha_2-1)$$
$$+ \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_2\alpha_3 + \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_5$$
$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_6 + \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_2\alpha_4$$
$$+ \tfrac{1}{12}(\alpha_1-1)(\alpha_1-2)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$+ \tfrac{1}{4}(\alpha_1-1)(\alpha_1-2)\alpha_3(\alpha_3-1) - (\alpha_1-1)\alpha_2\alpha_5$$
$$- \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3 - (\alpha_1-1)\alpha_3\alpha_4 - (\alpha_1-1)\alpha_7$$
$$- \tfrac{1}{24}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-3) - \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1)$$
$$- \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_4 - \alpha_2\alpha_6 - \alpha_3\alpha_5 - \tfrac{1}{2}\alpha_4(\alpha_4-1) - \alpha_8$$

$$D(n-8,6,2)_{(a)} = (1/2^5.3^2.7)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-5)(\alpha_1-8)(\alpha_1$$
$$+ \tfrac{1}{72}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1^2-15\alpha_1+53)\alpha_2$$
$$+ \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-5)(\alpha_1-7)\alpha_3$$
$$+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1^2-9\alpha_1+17)\alpha_2(\alpha_2-1)$$
$$+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_4$$
$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2\alpha_3 + \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3$$
$$- \tfrac{1}{4}(\alpha_1-1)(\alpha_1-2)\alpha_3(\alpha_3-1) - (\alpha_1-1)\alpha_3\alpha_4$$
$$- \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_6 - (\alpha_1-1)\alpha_7$$
$$+ \tfrac{1}{6}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-6) + \alpha_2(\alpha_2-2)\alpha_4$$
$$+ \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1) + (\alpha_2-1)\alpha_6$$

$$D(n-8,5,3)_{(a)} = (1/2.6!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-5)(\alpha_1-9)(\alpha_1$$
$$+ \tfrac{1}{72}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1^2-15\alpha_1+53)\alpha_2$$
$$+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1^2-9\alpha_1+17)\alpha_2(\alpha_2-1)$$
$$+ \tfrac{1}{120}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-9)\alpha_3$$
$$- \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_4$$
$$+ \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2\alpha_3 - \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_5$$
$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$+ \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_3(\alpha_3-1) + \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3$$
$$+ (\alpha_1-1)\alpha_3\alpha_4 - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_6$$
$$- \tfrac{1}{6}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-4) - \alpha_2(\alpha_2-2)\alpha_4$$
$$+ \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1) - \alpha_2\alpha_6 + \alpha_3\alpha_5$$

$$D(n-8,4^2)_{(a)} = (1/5!4!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-5)(\alpha_1-10)(\alpha_1$$
$$+ \tfrac{1}{360}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(2\alpha_1^2-30\alpha_1+103)\alpha_2$$
$$- \tfrac{1}{120}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1^2-12\alpha_1+47)\alpha_3 - \text{(see next p}$$

$$
\begin{aligned}
&- \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_4 \\
&+ \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_2(\alpha_2-1) \\
&+ \tfrac{1}{6}\alpha_1(\alpha_1^2-6\alpha_1+11)\alpha_2\alpha_3 - \tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_5 \\
&+ \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_3(\alpha_3-1) - (\alpha_1-1)\alpha_2\alpha_5 \\
&- \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3 + (\alpha_1-1)\alpha_3\alpha_4 \\
&+ \alpha_2(\alpha_2-2)\alpha_4 - (\alpha_2-1)\alpha_3(\alpha_3-1) \\
&+ \tfrac{1}{4}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-5) - \alpha_3\alpha_5 + \alpha_4(\alpha_4-1).
\end{aligned}
$$

3. The characters of the four-element-partition representations $D(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. On developing, as before, the four-rowed determinant $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ in terms of the first row we obtain

$$
\begin{aligned}
\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} &= q_{\lambda_1}(\boldsymbol{s})\{\lambda_2, \lambda_3, \lambda_4\} \\
&- q_{\lambda_1+1}(\boldsymbol{s})[\{\lambda_2, \lambda_3, \lambda_4-1\} + \{\lambda_2, \lambda_3-1, \lambda_4\} + \{\lambda_2-1, \lambda_3, \lambda_4\}] \\
&+ q_{\lambda_1+2}(\boldsymbol{s})[\{\lambda_2, \lambda_3-1, \lambda_4-1\} + \{\lambda_2-1, \lambda_3, \lambda_4-1\} + \{\lambda_2-1, \lambda_3-1, \lambda_4\}] \\
&- q_{\lambda_1+3}(\boldsymbol{s})\{\lambda_2-1, \lambda_3-1, \lambda_4-1\}.
\end{aligned}
$$

The coefficient of $-q_{\lambda_1+1}(\boldsymbol{s})$ in this expression is

$$\frac{\partial}{\partial s_1}\{\lambda_2, \lambda_3, \lambda_4\}$$

whilst that of $q_{\lambda_1+2}(\boldsymbol{s})$ is

$$\tfrac{1}{2}\left(\frac{\partial^2}{\partial s_1^2} - 2\frac{\partial}{\partial s_2}\right)\{\lambda_2, \lambda_3, \lambda_4\}.$$

Similarly the coefficient of $-q_{\lambda_1+3}(\boldsymbol{s})$ is

$$\tfrac{1}{6}\left\{\frac{\partial^2}{\partial s_1^3} - 3\frac{\partial}{\partial s_1}\left(2\frac{\partial}{\partial s_2}\right) + 2\frac{\partial}{\partial s_1}\left(3\frac{\partial}{\partial s_3}\right)\right\}\{\lambda_2, \lambda_3, \lambda_4\}$$

but it is simpler to leave it as it is. On equating the coefficients of

$$\frac{1}{\alpha_1!\cdots\alpha_n!}\left(\frac{s_1}{1}\right)^{\alpha_1}\cdots\left(\frac{s_n}{n}\right)^{\alpha_n}$$

on both sides we obtain

$$
\begin{aligned}
D(\lambda_1, \lambda_2, \lambda_3, \lambda_4)_{(\alpha)} &= \sum_{(\beta)}\{\lambda_2, \lambda_3, \lambda_4\}_{(\beta)}\binom{\alpha_1}{\beta_1}\cdots\binom{\alpha_n}{\beta_n} \\
&- \sum_{(\beta')}\{\lambda_2, \lambda_3, \lambda_4\}_{(\beta)}\binom{\alpha_1}{\beta_1'}\cdots\binom{\alpha_n}{\beta_n} \\
&+ \tfrac{1}{2}\sum_{(\beta'')}[\{\lambda_2, \lambda_3, \lambda_4\}_{\beta_1''+2} - \{\lambda_2, \lambda_3, \lambda_4\}_{\beta_2''+1}]\binom{\alpha_1}{\beta_1''}\binom{\alpha_2}{\beta_2''}\cdots\binom{\alpha_n}{\beta_n} \\
&- \sum_{(\gamma)}\{\lambda_2-1, \lambda_3-1, \lambda_4-1\}_{(\gamma)}\binom{\alpha_1}{\gamma_1}\cdots\binom{\alpha_n}{\gamma_n}.
\end{aligned}
$$

In this formula (β) is an arbitrary class of the symmetric group on $\lambda_2+\lambda_3+\lambda_4$ letters; (β') an arbitrary class of the symmetric group on $\lambda_2+\lambda_3+\lambda_4-1$ letters; (β'') an arbitrary class of the symmetric group on $\lambda_2+\lambda_3+\lambda_4-2$ letters and (γ) an arbitrary class of the symmetric group on $\lambda_2+\lambda_3+\lambda_4-3$ letters. In the second summation $\beta_1=\beta_1'+1$ and in the third summation the coefficient (in square parentheses) is the difference between the characters of $D(\lambda_2,\lambda_3,\lambda_4)$ for the classes $(\beta_1''+2,\beta_2'',\cdots,\beta_n)$ and $(\beta_1'',\beta_2''+1,\cdots,\beta_n)$ of the symmetric group on $\lambda_2+\lambda_3+\lambda_4$ letters. The first two summations may be combined in exactly the same manner as in the case of the three-element-partition representations and we thus obtain the formula:

$$
\begin{aligned}
(3)\quad & D(\lambda_1,\lambda_2,\lambda_3,\lambda_4)_{(a)} \\
&= \sum_{(\beta')}\frac{1}{\beta_1!}\{\lambda_2,\lambda_3,\lambda_4\}_{(\beta)}\alpha_1(\alpha_1-1)\cdots(\alpha_1-\beta_1'+1)(\alpha_1-2\beta_1'-1)\binom{\alpha_2}{\beta_2}\cdots\binom{\alpha_n}{\beta_n} \\
&+ \sum_{(\delta)}\{\lambda_2,\lambda_3,\lambda_4\}_{(\delta)}\binom{\alpha_2}{\delta_2}\cdots\binom{\alpha_n}{\delta_n} \\
&+ \tfrac{1}{2}\sum_{(\beta')}\left[\{\lambda_2,\lambda_3,\lambda_4\}_{\beta_1''+2}-\{\lambda_2,\lambda_3,\lambda_4\}_{\beta_2''+1}\right]\binom{\alpha_1}{\beta_1''}\binom{\alpha_2}{\beta_2''}\cdots\binom{\alpha_n}{\beta_n} \\
&- \sum_{(\gamma)}\{\lambda_2-1,\lambda_3-1,\lambda_4-1\}_{(\gamma)}\binom{\alpha_1}{\gamma_1}\cdots\binom{\alpha_n}{\gamma_n}
\end{aligned}
$$

where (δ) is any class of the symmetric group on $\lambda_2+\lambda_3+\lambda_4$ letters *which does not contain any unary cycle* (= fixed letter).

As examples of the use of this formula we list the following:

$$
\begin{aligned}
D(n-3,1^3)_{(a)} &= \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)-(\alpha_1-1)\alpha_2+\alpha_3 \\
D(n-4,2,1^2)_{(a)} &= \tfrac{1}{8}\alpha_1(\alpha_1-2)(\alpha_1-3)(\alpha_1-5) \\
&\quad -\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2-\tfrac{1}{2}\alpha_2(\alpha_2-1)+\alpha_4 \\
D(n-5,3,1^2)_{(a)} &= \tfrac{1}{20}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7) \\
&\quad -(\alpha_1-1)\alpha_2(\alpha_2-2)+\alpha_5 \\
D(n-5,2^2,1)_{(a)} &= \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-5)(\alpha_1-6) \\
&\quad -\tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_2 \\
&\quad -\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_3+\tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1) \\
&\quad +(\alpha_1-1)\alpha_4-\alpha_2\alpha_3 \\
D(n-6,4,1^2)_{(a)} &= \tfrac{1}{72}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-5)(\alpha_1-9) \\
&\quad +\tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_2+\alpha_1(\alpha_1-2)\alpha_2 \\
&\quad +\tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_3+(\alpha_1-1)\alpha_3 \\
&\quad -\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)-(\alpha_1-1)\alpha_2\alpha_3 \\
&\quad -\tfrac{1}{3}\alpha_2(\alpha_2-1)(\alpha_2-2)+\tfrac{1}{2}\alpha_3(\alpha_3-1)+\alpha_6
\end{aligned}
$$

$$D(n-6,3,2,1)_{(a)} = \tfrac{1}{45}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-6)(\alpha_1-8)$$
$$-\tfrac{1}{3}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_3-(\alpha_1-1)\alpha_3$$
$$-\alpha_3(\alpha_3-1)+(\alpha_1-1)\alpha_5$$

$$D(n-6,2^3)_{(a)} = \tfrac{1}{144}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-5)(\alpha_1-6)(\alpha_1-7)$$
$$-\tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_2$$
$$-\tfrac{1}{2}\alpha_1(\alpha_1-1)\alpha_2-\tfrac{1}{2}\alpha_1\alpha_2(\alpha_2-3)$$
$$-\tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_3$$
$$+\tfrac{1}{4}(\alpha_1+1)(\alpha_1-2)\alpha_2(\alpha_2-1)$$
$$+\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_4-(\alpha_1-1)\alpha_2\alpha_3$$
$$+\tfrac{1}{2}\alpha_2(\alpha_2-1)(\alpha_2-2)-(\alpha_2-1)\alpha_4+\alpha_3(\alpha_3-2)$$

$$D(n-7,5,1^2)_{(a)} = \tfrac{1}{336}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)(\alpha_1-6)(\alpha_1-11)$$
$$+\tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-5)(\alpha_1-6)\alpha_2$$
$$+\tfrac{1}{8}\alpha_1(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)\alpha_3$$
$$+\tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_4$$
$$-\tfrac{1}{12}(\alpha_1+1)(\alpha_1-1)(\alpha_1-6)\alpha_2(\alpha_2-1)$$
$$-\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_3-(\alpha_1-1)\alpha_2\alpha_4$$
$$-\tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$-\tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3+\alpha_3\alpha_4+\alpha_7$$

$$D(n-7,4,2,1)_{(a)} = \tfrac{1}{144}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)(\alpha_1-7)(\alpha_1-10)$$
$$+\tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-5)(\alpha_1-6)\alpha_2$$
$$-\tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_3$$
$$-\tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_4$$
$$-\tfrac{1}{12}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_2(\alpha_2-1)$$
$$-\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_3+(\alpha_1-1)\alpha_2\alpha_4$$
$$+\tfrac{1}{6}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$-\tfrac{1}{2}(\alpha_1-1)\alpha_3(\alpha_3-1)+(\alpha_1-1)\alpha_6$$
$$-\tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3-\alpha_3\alpha_4$$

$$D(n-7,3^2,1)_{(a)} = \tfrac{1}{240}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)(\alpha_1-8)(\alpha_1-9)$$
$$+\tfrac{1}{120}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-9)\alpha_2$$
$$-\tfrac{1}{8}\alpha_1(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)\alpha_3$$
$$-\tfrac{1}{6}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_4$$
$$+\tfrac{1}{12}(\alpha_1^2-5\alpha_1+12)(\alpha_1-1)\alpha_2(\alpha_2-1)$$
$$+\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_3$$
$$+\tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_5-(\alpha_1-1)\alpha_2\alpha_4$$
$$-\tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2)+\alpha_2\alpha_5$$
$$+\tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3-\alpha_3\alpha_4$$

$$D(n-7,3,2^2)_{(a)} = \tfrac{1}{240}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-6)(\alpha_1-7)(\alpha_1-9)$$
$$-\tfrac{1}{120}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-9)\alpha_2$$
$$-\tfrac{1}{8}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-6)\alpha_3+$$ (see next page)

$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_4$$
$$+ \tfrac{1}{12}\alpha_1(\alpha_1-1)(\alpha_1-5)\alpha_2(\alpha_2-3)$$
$$- \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_2\alpha_3$$
$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_5 - (\alpha_1-1)\alpha_2\alpha_4$$
$$+ \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-3)$$
$$+ \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_3 - \alpha_2\alpha_5 + \alpha_3\alpha_4.$$

$$D(n-8,6,1^2)_{(a)} = (1/2^4.5!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-6)(\alpha_1-7)(\alpha_1-$$
$$+ \tfrac{1}{240}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(3\alpha_1^2-42\alpha_1+139)\alpha_2$$
$$+ \tfrac{1}{20}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)\alpha_3$$
$$+ \tfrac{1}{48}\alpha_1(\alpha_1-2)(\alpha_1^2-8\alpha_1+31)\alpha_2(\alpha_2-1)$$
$$+ (\alpha_1-1)\alpha_2\alpha_3 - \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$- (\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3 - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2\alpha_4$$
$$- \tfrac{1}{8}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-3) - (\alpha_1-1)\alpha_2\alpha_5$$
$$- \tfrac{1}{2}\alpha_2(\alpha_2-1)\alpha_4 + \tfrac{1}{8}\alpha_1(\alpha_1-2)(\alpha_1-3)(\alpha_1-5)\alpha_4$$
$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_5 + \alpha_3\alpha_5$$
$$+ \tfrac{1}{2}\alpha_4(\alpha_4-1) + \alpha_8$$

$$D(n-8,5,2,1)_{(a)} = \tfrac{1}{630}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-6)(\alpha_1-8)(\alpha_1-12)$$
$$+ \tfrac{1}{45}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-6)(\alpha_1-8)\alpha_2$$
$$+ \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1^2-11\alpha_1+33)\alpha_3$$
$$- \tfrac{1}{3}(\alpha_1-1)(\alpha_1^2-5\alpha_1+3)\alpha_2\alpha_3$$
$$- \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_3(\alpha_3-1)$$
$$- \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_5 + (\alpha_1-1)\alpha_2\alpha_5$$
$$+ (\alpha_1-1)\alpha_{\cdot} - \alpha_2\alpha_3(\alpha_3-1) - \alpha_3\alpha_5$$

$$D(n-8,4,3,1)_{(a)} = \tfrac{1}{576}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-6)(\alpha_1-9)(\alpha_1-11)$$
$$+ \tfrac{1}{72}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-5)(\alpha_1-9)\alpha_2$$
$$- \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-5)(\alpha_1-6)\alpha_3$$
$$+ \tfrac{1}{24}\alpha_1(\alpha_1-2)(\alpha_1^2-8\alpha_1+19)\alpha_2(\alpha_2-1)$$
$$- \tfrac{1}{6}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_2\alpha_3$$
$$- \tfrac{1}{2}(\alpha_1-1)\alpha_2(\alpha_2-1)\alpha_3 + \tfrac{1}{4}\alpha_1(\alpha_1-3)\alpha_3(\alpha_3-1)$$
$$- \tfrac{1}{6}\alpha_1(\alpha_1-2)(\alpha_1^2-8\alpha_1+13)\alpha_4 + \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_6$$
$$- (\alpha_1-1)\alpha_3\alpha_4 - \tfrac{1}{12}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-3)$$
$$+ \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1) + \alpha_2\alpha_6 - \alpha_4(\alpha_4-1)$$

$$D(n-8,4,2^2)_{(a)} = (1/6!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)(\alpha_1-8)(\alpha_1-1$$
$$+ \tfrac{1}{180}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1^2-15\alpha_1+59)\alpha_2$$
$$- \tfrac{1}{30}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-4)(\alpha_1-8)\alpha_3$$
$$+ \tfrac{1}{3}(\alpha_1^2-3\alpha_1-1)\alpha_2(\alpha_2-1)(\alpha_2-2)$$
$$- \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_2\alpha_3 - \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_2(\alpha_2-1)$$
$$- \tfrac{1}{4}(\alpha_1-1)(\alpha_1-2)\alpha_3(\alpha_3-1) + \text{(see next page)}$$

$$+ \tfrac{1}{6}(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)\alpha_5$$
$$+ \tfrac{1}{2}(\alpha_1-1)(\alpha_1-2)\alpha_6 - (\alpha_1-1)\alpha_2\alpha_5$$
$$+ \tfrac{1}{3}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-3)$$
$$+ \tfrac{1}{2}\alpha_2\alpha_3(\alpha_3-1) + \alpha_3\alpha_5 - \alpha_2\alpha_6$$

$$(n-8, 3^2, 2)_{(\alpha)} = (1/8.5!)\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)(\alpha_1-9)(\alpha_1-10)$$
$$- \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-2)(\alpha_1-7)\alpha_2$$
$$- \tfrac{1}{20}\alpha_1(\alpha_1-1)(\alpha_1-3)(\alpha_1-4)(\alpha_1-7)\alpha_3$$
$$+ \tfrac{1}{24}\alpha_1(\alpha_1-1)(\alpha_1-4)(\alpha_1-5)\alpha_2(\alpha_2-1)$$
$$+ \tfrac{1}{3}\alpha_1(\alpha_1-2)(\alpha_1-4)\alpha_5 - (\alpha_1^2-3\alpha_1+1)\alpha_2\alpha_4$$
$$+ \tfrac{1}{2}\alpha_1(\alpha_1-3)\alpha_4$$
$$- \tfrac{1}{2}\alpha_1\alpha_2(\alpha_2-1) + (\alpha_1-1)\alpha_2(\alpha_2-2)\alpha_3$$
$$- \tfrac{1}{4}\alpha_2(\alpha_2-1)(\alpha_2-2)(\alpha_2-5) - \alpha_3\alpha_5 + \alpha_4(\alpha_4-1).$$

4. Concluding remarks. Whilst the expressions given in the preceding paragraphs are somewhat lengthy their evaluation for any given class (α) of the symmetric group on n letters is quite easy since a great many terms of each expression vanish for any given class (α). The formulae explicitly written out suffice to give all characters for $n \leqq 8$; for $n = 9$ they give all save the characters of the self-associated five-element-partition $(5, 1^4)$; it being always understood that of two associated partitions of n we apply our formula to the one whose partition contains the smaller number of elements. For $n = 10$ our formulae yield all characters save those corresponding to the partitions $(6, 1^4)$, $(5, 2, 1^3)$ the second of which is self associated. For $n = 11$ the only partitions not cared for are $(7, 1^4)$, $(6, 2, 1^3)$, $(6, 1^5)$, $(5, 3, 1^3)$ whilst for $n = 12$ they are $(8, 1^4)$, $(7, 2, 1^3)$, $(7, 1^5)$, $(6, 3, 1^3)$, $(6, 2^2, 1^2)$, $(6, 2, 1^4)$, $(5, 4, 1^3)$, $(5, 3, 2, 1^2)$. Thus if we are concerned only with partitions containing not more than four elements the characters are all cared for, up to $n = 12$, by our formulae. The missing partitions, up to $n = 12$, may be cared for since they all end in 1 by constructing formulae as described in the following paragraph.

Since

$$\{\lambda_1, \cdots, \lambda_k\} \cdot \{1\} = \{\lambda_1, \cdots, \lambda_k, 1\} + \{\lambda_1 + 1, \cdots, \lambda_k\} + \cdots + \{\lambda_1, \cdots, \lambda_k + 1\}$$

(p. 480) we find at once, on comparing the coefficients of

$$\frac{1}{\alpha_1! \cdots \alpha_n!}\left(\frac{s_1}{1}\right)^{\alpha_1} \cdots \left(\frac{s_n}{n}\right)^{\alpha_n}$$

(where (α) is an arbitrary class of the symmetric group on $\lambda_1 + \cdots + \lambda_k + 1$ letters) on both sides the formula:

$$(4) \qquad D(\lambda_1, \cdots, \lambda_k, 1) = \alpha_1 D(\lambda_1, \cdots, \lambda_k)_{(\alpha')} - [D(\lambda_1 + 1, \cdots, \lambda_k)_{(\alpha)} + \cdots + D(\lambda_1, \cdots, \lambda_k + 1)_{(\alpha)}]$$

where $\alpha'_1 = \alpha_1 - 1$ so that $(\alpha') = (\alpha_1 - 1, \alpha_2, \cdots, \alpha_n)$ is a class of the symmetric group on $\lambda_1 + \lambda_2 + \cdots + \lambda_k$ letters. As an instance of the application of this formula consider $D(n-4, 1^4)$: here

$$\lambda_1 = n - 4, \quad \lambda_2 = \lambda_3 = \lambda_4 = 1$$

and so

$$D(n-4, 1^4)_{(\alpha)} = \alpha_1 D(n-4, 1^3)_{(\alpha')} - [D(n-3, 1^3)_{(\alpha)} + D(n-4, 2, 1^2)_{(\alpha)}].$$

Since n does not enter explicitly the previously derived formula for $D(n-3, 1^3)_{(\alpha)}$ we read off the expression for $D(n-4, 1^3)_{(\alpha')}$ by merely replacing α_1 by $\alpha_1 - 1$ in the expression for $D(n-3, 1^3)_{(\alpha)}$. In this way we find

$$\begin{aligned} D(n-4, 1^4)_{(\alpha)} &= \tfrac{1}{24}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4) \\ &\quad - \tfrac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_2 + (\alpha_1 - 1)\alpha_3 \\ &\quad + \tfrac{1}{2}\alpha_2(\alpha_2 - 1) - \alpha_4 \end{aligned}$$

Similarly

$$\begin{aligned} D(n-5, 2, 1^3)_{(\alpha)} &= \tfrac{1}{30}\alpha_1(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 6) \\ &\quad - \tfrac{1}{3}\alpha_1(\alpha_1 - 2)(\alpha_1 - 4)\alpha_2 \\ &\quad + \tfrac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_3 + \alpha_2\alpha_3 - \alpha_5 \\ D(n-5, 1^5)_{(\alpha)} &= \tfrac{1}{120}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 5) \\ &\quad - \tfrac{1}{6}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)\alpha_2 \\ &\quad + \tfrac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_3 \\ &\quad + \tfrac{1}{2}(\alpha_1 - 1)\alpha_2(\alpha_2 - 1) - (\alpha_1 - 1)\alpha_4 - \alpha_2\alpha_3 + \alpha_5 \\ D(n-6, 3, 1^3)_{(\alpha)} &= \tfrac{1}{72}\alpha_1(\alpha_1 - 1)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 5)(\alpha_1 - 8) \\ &\quad - \tfrac{1}{12}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)\alpha_2 \\ &\quad + (\alpha_1 - 1)(\alpha_1 - 3)\alpha_2 \\ &\quad - \tfrac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_2(\alpha_2 - 1) \\ &\quad + \tfrac{1}{3}\alpha_2(\alpha_2 - 1)(\alpha_2 - 2) \\ &\quad + \tfrac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_3 + (\alpha_1 - 1)\alpha_2\alpha_3 \\ &\quad + \tfrac{1}{2}\alpha_3(\alpha_3 - 1) - \alpha_6 \end{aligned}$$

These formulae, together with those already given, furnish all the characters of the various symmetric groups for $n \leq 11$. The characters of the symmetric groups up to $n = 13$ have recently been published (**4**) and the provision of the formulae of the present paper would be without point if they

merely enabled us to provide data already known by other methods. The whole point of the present investigation is that the formulae given are available for *any value* of n. To be explicit suppose we wish to know the characters of the symmetric group on $n = 16$ letters. This group has 231 irreducible representations and the mere printing of the character table presents difficulties. Of the 231 representations some 88 $(= 44 \times 2)$ have their characters given by the formulae of the present paper. In addition to these we have the identity representation and its associate, the alternating representation, so that we know, without recurrence methods, the characters of 90 out of the 231 irreducible representations. Of the 231 partitions of 16, sixty-four contain not more than four elements and our formulae furnish the characters of 41 of the 64 corresponding irreducible representations.

THE JOHNS HOPKINS UNIVERSITY,
INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

1. F. D. Murnaghan, "On the representations of the symmetric group," *American Journal of Mathematics*, **59** (1937), pp. 437-488.
2. G. Frobenius, "Über die Charaktere der symmetrischen Gruppe," *Berliner Berichte* (1900), pp. 516-534.
3. D. E. Littlewood and A. R. Richardson, "Some special S-functions and q-series," *Quarterly Journal of Mathematics*, Oxford Ser. **6** (1935), pp. 184-198.
4. M. Zia-ud-Din, "The characters of the symmetric group of degrees 12 and 13," *Proceedings of the London Mathematical Society* (2), **42** (1937), pp. 340-355.

ISOTROPIC STATIC SOLUTIONS OF THE FIELD EQUATIONS IN EINSTEIN'S THEORY OF GRAVITATION.*

By P. Y. Chou.

1. Introduction. In the classic Schwarzschild's solution of the gravitational field of a mass particle it is well known that a system of isotropic coördinates exists in which the velocity of light is the same in every direction at any point exterior to the particle. The question naturally arises whether there are other isotropic solutions of Einstein's field equations, and if they exist, we should like to know the kind of distribution of matter that produces the corresponding gravitational fields.

We shall limit our investigation to static fields only and shall give a detailed proof of the necessary and sufficient condition for the existence of isotropic fields in empty space, while proof for the corresponding theorem within matter will only be indicated in general outlines. We shall see that for a type of problem involving the determination of the field of a single body so that space outside the body is free from other singularities, there are two classes of solutions and as special cases of each class we find respectively Schwarzschild's solution and the solution for a semi-infinite plane with variable distribution of mass which was not known before. As a third example of our general result we give Kasner's solution which forms a class by itself. From the field equations within matter in general we can also prove incidentally that the Einstein static universe is the only solution for a closed static space filled with matter which is kept at constant pressure everywhere without assuming the spherical symmmetry property of the universe to start with.

2. Isotropic static gravitational fields in empty space. The arc element of a static four dimensional continuum takes the form

$$ds^2 = U^2 dl^2 + g_{ij} dx^i dx^j, \qquad (i, j = 1, 2, 3) \tag{2.1}$$

* Received April 23, 1937.

where $t = x^0$ denotes the time-like and x^1, x^2, x^3, the space-like coördinates and the functions U, g_{ij} are functions of x^1, x^2, x^3 only. We use Latin letters to denote the indices, 1, 2, 3, of the space variables, while Greek letters can take four indices including the time coördinate x^0. The ten field equations with the cosmological constant absent,

$$G_{\alpha\beta} = 0 \tag{2.2}$$

are split in the present case into

$$G_{ij} = R_{ij} + \frac{1}{U} U_{i,j} = 0, \qquad G_{0i} = 0, \qquad G_{00} = UU^i{}_{,i} = 0, \tag{2.3}$$

where R_{ij} is the contracted Riemann-Christoffel tensor for the 3-space, $t = \text{const.}$, of the arc-element (2.1), and we denote, for brevity,

$$U_i = \frac{\partial U}{\partial x^i}, \qquad U^i = g^{ij} U_j;$$

$$U_{i,j} = \frac{\partial U_i}{\partial x^j} - \left\{ \begin{matrix} h \\ ij \end{matrix} \right\} U_h, \qquad U^i{}_{,i} = g^{ij} U_{i,j},$$

covariant differentiation being taken with respect to the metric tensor g_{ij}. Since we only deal with transformations of the space-like coördinates, U is an invariant, U_i and its higher covariant derivatives are all tensors.

We assume that within the regions of the 3-space under investigation the functions U and g_{ij} possess continuous derivatives to any degree we need. Furthermore we are only interested in the real solutions of (2.3) for these gravitational potentials and real transformations of coördinates so that two solutions which can not be transformed into each other by a real transformation are considered as distinct. Next we shall establish the validity of the following theorem:

THEOREM. *A necessary and sufficient condition for the static field equations* (2.3) *to possess isotropic solutions in which the arc-element* (2.1) *can be transformed into the canonical form*

$$ds^2 = U^2 dt^2 - e^{-2\sigma}(dx^2 + dy^2 + dz^2) \tag{2.4}$$

is that the function U should satisfy the following set of differential equations,

$$U_{i,j} = \frac{6U}{c+U^2}\,(U_iU_j - \tfrac{1}{3}g_{ij}U^hU_h), \quad \text{and} \quad R = 0, \tag{2.5}$$

where R is the scalar curvature invariant for the 3-space, $t = const.$, and c is an arbitrary constant.

If we contract (2.5) with U^j, the resultant equations are integrable giving

$$U^hU_h = -k^2(c+U^2)^4 \tag{2.6}$$

where k^2 is another constant of integration.

We begin to prove the theorem by establishing the necessary condition first. If the arc-element of the 3-space in (2.1) is conformal to a flat space as in (2.4), then the contracted Riemann-Christoffel tensor R_{ij} satisfies the necessary condition [1]

$$R_{ij,k} - R_{ik,j} - \tfrac{1}{4}(g_{ij}R_{,k} - g_{ik}R_{,j}) = 0 \tag{2.7}$$

where $R_{ij,k}$ denotes as usual the covariant derivative of R_{ij} with respect to the metric tensor g_{ij} in (2.1). From (2.3) we have $R = 0$ and

$$U^hR_{hijk} - \frac{1}{U}\,(U_{i,j}U_k - U_{i,k}U_j) = 0 \tag{2.8}$$

in which we have utilized the well-known relation

$$U_{i,jk} - U_{i,kj} = U^hR_{hijk}. \tag{2.9}$$

Now for any 3-space Weyl's conformal tensor vanishes identically or the Riemann-Christoffel tensor is expressible as [2]

$$R_{hijk} = -g_{jh}R_{ik} + g_{hk}R_{ij} - g_{ik}R_{hj} + g_{ij}R_{hk} - \frac{R}{2}\,(g_{hk}g_{ij} - g_{hj}g_{ik}). \tag{2.10}$$

Combining (2.3) with (2.8) and (2.10) we obtain the following relation on the covariant derivatives of U

$$2(U_{i,j}U_k - U_{i,k}U_j) + g_{ij}U^hU_{h,k} - g_{ik}U^hU_{h,j} = 0. \tag{2.11}$$

---

[1] L. P. Eisenhart, *Riemannian Geometry* (1926), which will be referred to as "R. G.", p. 92.

[2] R. G., p. 90.

To solve (2. 11) for U_{ij} in terms of U and its first derivatives we contract (2. 11) by U^k and U^i in succession and eliminate the intermediate expression, $U^kU_{k,j}$. The result can be written

$$U_{i,j} = \alpha(U_iU_j - \tfrac{1}{3}g_{ij}U^hU_h) \tag{2. 12}$$

where α, an invariant, is defined as $\alpha = 3U^hU_{h,k}U^k/2(U_lU^l)^2$.

This function α can be determined in the following way. We regard (2. 12) as a set of equations defining the function $U_{i,j}$ which, on the other hand, should satisfy the condition of integrability (2. 9). Putting (2. 12) in (2. 9), then contracting with U^i and meanwhile utilizing the relations (2. 10) and the field equations (2. 3), we find that α must be a function of U alone and furthermore its derivative with respect to U is given by

$$\frac{d\alpha}{dU} = \left(\frac{\alpha}{U} - \frac{\alpha^2}{3}\right) \tag{2. 13}$$

Integrating, we get

$$\alpha = 6U/(c + U^2) \tag{2. 14}$$

where c is a constant of integration. Hence (2. 5) as a necessary condition for the existence of isotropic solutions of the field equations (2. 3) is established.

We shall show conversely that if this condition holds true, then (2. 3) are satisfied and from the way of deriving (2. 5), condition (2. 7) is satisfied which is also sufficient for the 3-space in (2. 1) to be conformal to a flat space.

Differentiate (2. 5) covariantly with respect to g_{ij} and form the left-hand side of (2. 9). Inserting the expression for R_{hijk} in the right-hand side from (2. 10), we find

$$\begin{aligned}(2.15)\quad &\frac{2}{c + U^2}U^hU_h(g_{ij}U_k - g_{ik}U_j)\\ &= U_jR_{ik} - U_kR_{ij} + g_{ik}U^hR_{hj} - g_{ij}U^hR_{hk} + \frac{R}{2}(g_{ij}U_k - g_{ik}U_j).\end{aligned}$$

Contract the indices i and k by g^{ik} and obtain

$$U^kR_{kj} = -\frac{4}{c + U^2}U^hU_hU_j. \tag{2. 16}$$

Contracting (2. 15) with U^k and eliminating U^kR_{kj} by (2. 16) we get finally

$$(2.17)\quad R_{ij} = -\frac{6}{c+U^2}(U_iU_j - \tfrac{1}{3}g_{ij}U^hU_h) + \frac{R}{2}\left(g_{ij} - \frac{U_iU_j}{U^hU_h}\right).$$

The function R in (2.17) can be set equal to zero as follows: By utilizing the Bianchi relation which holds for any Riemannian space,[3]

$$(2.18)\qquad R^i{}_{j,i} - \tfrac{1}{2}R_{,j} = 0,$$

we can form out of (2.5), (2.6) and (2.17) a differential equation in R which can be integrated. If we choose the constant of integration properly R vanishes identically. The fact that R can be different from zero points out the possibility of existence of isotropic solutions of the field equations within matter as we shall see below.

3. Field of a single body; Kasner and Schwarzschild's solutions and field of a semi-infinite plane. We have seen in the last section that the existence and finding of the isotropic gravitational fields are equivalent to the solution of the equations (2.5). For the general problem which may involve a number of bodies, the solutions of (2.5) are very difficult to ascertain. But for the type of problem which contains only one body, then the space outside the body is free from other singularities and we can make a transformation of coördinates such that in this new system of reference the function U becomes a function of the coördinate u only. Moreover we limit ourselves to the consideration of the class of problems in which with the coördinate u thus specified the spatial part of the arc-element (2.1) can still assume the orthogonal form. Hence (2.1) can be written as

$$(3.1)\qquad ds^2 = U^2(u)dl^2 + g_{11}du^2 + g_{22}dv^2 + g_{33}dw^2,$$

where g_{11}, g_{22}, g_{33} are unknown functions of u, v, w to be determined.

The equations (2.5) can be computed for the quadratic form (3.1). They are immediately integrable giving

$$\frac{dU}{du} = k(c+U^2)^2(-g_{11})^{\frac{1}{2}},$$

$$(3.2)\quad g_{11} = g_1(u),\quad g_{22} = g_2(v,w)(c+U^2)^{-2},\quad g_{33} = g_3(v,w)(c+U^2)^{-2},$$

[3] R. G., p. 82.

where k is a constant of integration and g_1, g_2, g_3 are arbitrary functions of their arguments indicated. We notice that equations (3.2) also satisfy (2.6) as they should.

Since any 2-space is conformal to the flat 2-space and g_{11} is a function of u only, we may write (3.1) without any loss of generality in the following form

$$ds^2 = U^2(u)dt^2 - \frac{1}{(c+U^2)^2}[du^2 + e^{\chi(v,w)}(dv^2+dw^2)], \tag{3.3}$$

where

$$\frac{dU}{du} = k(c+U^2).$$

So far only the equations (2.5) involving U are satisfied. In order to satisfy every member of the field equations (2.3), the scalar curvature invariant R of the 3-space $t = \text{const.}$ in (3.3) must vanish. This condition leads to the following differential equation in χ [4]

$$\frac{\partial^2\chi}{\partial v^2} + \frac{\partial^2\chi}{\partial w^2} - 8k^2ce^{\chi} = 0. \tag{3.4}$$

The constant c may be either negative, zero, or positive and can be normalized to be -1, 0, $+1$ respectively for simplicity. The three values of c correspond to three distinct classes of solutions. We shall show first that for the class $c = 0$, there is only one solution in the class which was discovered by Kasner and all other members of the class are transformable into Kasner's solution.

When $c = 0$, (3.4) becomes Laplace's equation for χ in the vw-plane. If we take the solution $\chi = 0$, then (3.3) becomes Kasner's solution [5]

$$ds^2 = (ku)^{-2}dt^2 - k^4u^4(du^2+dv^2+dw^2) \tag{3.5}$$

which can be considered as describing the field of an infinite plane of constant surface density coinciding with the vw-plane. But this is also the only solution of the class. For if we take the general real solution of χ to be

---

[4] The computation of $R = 0$ can be accomplished by setting $T_4^4 = 0$ on p. 256 in Tolman's *Relativity, Thermodynamics, and Cosmology* (1934).

[5] E. Kasner, *Transactions of the American Mathematical Society*, vol. 27 (1925), p. 160.

$$\chi = f(v + iw) + f(v - iw) = f + f^*$$

where f is an arbitrary analytic function of the complex variable $z = v + iw$, then

$$e^{\chi}(dv^2 + dw^2) = e^f dz e^{f^*} dz^* = dZ dZ^* = dV^2 + dW^2,$$

where

$$dZ = e^f dz = dV + i dW.$$

For the other two classes of solutions $c = \pm 1$, there is an infinite number in each class. This is due to the fact that to each solution of the partial differential equation (3. 4) we can make a transformation of the coördinates from the (u, v, w)-system to the canonical system (x, y, z) of (2. 4). In these transformations all of the three variables u, v, w are involved in general and the final functions $\sigma(x, y, z)$ and $U(x, y, z)$ are necessarily distinct if we start with the distinct function χ. The converse argument also holds. If two solutions $U_1(x, y, z)$, $\sigma_1(x, y, z)$ and $U_2(x, y, z)$, $\sigma_2(x, y, z)$ in the (x, y, z)-system of (2. 4) can be transformed into each other by coördinate transformation, then in the (u, v, w)-system they should have the same function $\chi(v, w)$.

Due to the non-linear character of the equation (3. 4), its general solution is difficult to obtain. But we can give two special solutions, one in each class. Consider the case $c = -1$. Then from (3. 3) we can take

$$U^2 = \tan h^2 ku. \tag{3. 6}$$

As a special solution of (3. 4) we assume χ to be a function of v only. Moreover from this ordinary differential equation in v we choose the solution

$$e^{\chi} = \sin^2 \theta/4k^2, \quad \text{where} \quad dv = d\theta/\sin\theta. \tag{3. 7}$$

If we put $e^{-2ku} = m/2r$, $k^2 = 1/16m^2$, the arc-element (3. 3) is recognizable as Schwarzschild's solution in isotropic coördinates.

As a solution for the second class $c = +1$, we also take $\chi = \chi(v)$ and the function

$$e^{\chi} = 1/4k^2v^2. \tag{3. 8}$$

Take $U^2 = \cot^2 ku$ and by normalizing the coördinates such that $2ku = \phi$, $2kv = \rho$, and $2kw = z$, we find

$$ds^2 = \cot^2(\phi/2)\, dt^2 - \frac{\sin^4 \phi/2}{(2k)^2\rho^2}\,(d\rho^2 + \rho^2 d\phi^2 + dz^2). \tag{3. 9}$$

The geometrical significance of the coördinates ρ, ϕ, z is self-evident. The gravitational potentials in (3. 9) are even functions of ϕ periodic in ϕ with the period 2π, their only singularity being on the positive x-axis $\phi = 0$. If we compare (3. 9) and (3. 5), we notice that at points where ϕ is small and ρ large the field defined by (3. 9) is very much the same as that of Kasner's solution, (3. 5). Hence we may interpret (3. 9) as the gravitational field of a semi-infinite plane coinciding with the positive half of the zx-plane with variable distribution of mass upon the plane.

4. Isotropic fields within matter. The static field equations for the arc-element (2. 1) within matter are

$$
\begin{aligned}
R_{ij} &= -\frac{U}{1} U_{i,j} + 4\pi\left(\rho_{00} - p_0 + \frac{\Lambda}{4\pi}\right) g_{ij}, \\
U^i{}_{,i} &= -4\pi\left(\rho_{00} + 3p_0 - \frac{\Lambda}{4\pi}\right) U, \qquad (4.1) \\
R &= 16\pi\left(\rho_{00} + \frac{\Lambda}{8\pi}\right),
\end{aligned}
$$

where Λ is the cosmological constant, ρ_{00} the proper density and p_0 the pressure, ρ_{00} and p_0 being invariant functions of the space-like coördinates and measured in relativistic units.

To the field equations (4. 1) we can add the dynamical equations of motion,

$$P^{\mu\nu}{}_{,\nu} = 0; \qquad P^{\mu\nu} = (\rho_{00} + p_0) v^\mu v^\nu - p_0 g^{\mu\nu}.$$

In the present static case since $v^i = 0$, $v^0 = 1/U$, we find

$$P^{ij} = -p_0 g^{ij}, \qquad P^{i0} = 0, \qquad P^{00} = \rho_{00}/U^2$$

and the equations of motion degenerate into

$$p_{0,i} = -\frac{1}{U}(\rho_{00} + p_0) U_i \qquad (4.2)$$

which can also be derived by inserting (4. 1) in the Bianchi identity (2. 18).

From (4. 1) and (4. 2) we can immediately draw the conclusion that if matter is kept at a constant non-negative pressure everywhere in a closed space ($\Lambda > 0$), then the only solution of (4. 1) is the Einstein static universe of relativistic cosmology. For in such a space U must be constant by (4. 2) and from (4. 1) it follows that

$$(4.3) \qquad \rho_{00} + 3p_0 = \frac{\Lambda}{4\pi}, \quad \text{and} \quad R_{ij} = 4\pi\left(\rho_{00} - p_0 + \frac{\Lambda}{4\pi}\right) g_{ij}$$

which is the necessary and sufficient condition for a 3-space to possess constant Riemannian curvature.[6] It is also well known that an isotropic system of coördinates exists for the Einstein static universe.

In the general case when p_0 is not constant, a necessary and sufficient condition for (4.1) to admit isotropic solutions is still (2.7). We can substitute R_{ij} and R from (4.1) into (2.7) and take into account relations (2.9) and (2.10). Finally we obtain an equation involving the covariant derivatives of U analogous to (2.11). By a similar process which was used in deriving (2.12) we find

$$(4.4) \quad U_{i,j} = \alpha(U_iU_j - \tfrac{1}{3}g_{ij}U^hU_h) - 2\pi\left(\rho_{00} + 3p_0 - \frac{\Lambda}{4\pi}\right)\left(g_{ij} - \frac{U_iU_j}{U^hU_h}\right)U,$$

where α plays the same rôle as α in (2.12).

The condition of integrability (2.9) for $U_{i,j}$ in (4.4) must be satisfied. We thus arrive at the conclusion that the functions α, ρ_{00}, U^hU_h must be all functions of U alone and the derivatives of α and ρ_{00} are related by

$$(4.5) \quad \tfrac{1}{3}\left(\frac{d\alpha}{dU} - \frac{\alpha}{U} + \frac{\alpha^2}{3}\right)U^hU_h - 2\pi(\rho_{00} + p_0) - \frac{8\pi}{3}\left(\rho_{00} + 3p_0 - \frac{\Lambda}{4\pi}\right)U\alpha$$
$$- 4\pi^2\left(\rho_{00} + 3p_0 - \frac{\Lambda}{4\pi}\right)^2\frac{U^2}{U^hU_h} + 2\pi U\frac{d\rho_{00}}{dU} = 0.$$

If we contract (4.4) by U^i, we obtain

$$(4.6) \qquad \frac{d}{dU}(U^hU_h) = \tfrac{4}{3}\alpha U^hU_h$$

which has the solution,

$$(4.7) \qquad U^hU_h = -k^2\exp(\tfrac{4}{3}\textstyle\int\alpha dU),$$

where k^2 is a constant of integration. If ρ_{00} and p_0 are related in the form of an equation of state as in the theory of gases, then equations (4.2), (4.5) and (4.7) can be solved completely for the functions ρ_{00}, p_0, α and U^hU_h in terms of U. The above results can be summarized in the following theorem:

---

[6] J. A. Schouten and D. J. Struik, *American Journal of Mathematics*, vol. 43 (1921), p. 214; or inserting R_{ij} from (4.3) in (2.10) we obtain the same result (R. G. 84).

THEOREM. *If in* (4.1) ρ_{00} *is a given function of* p_0, *then a necessary and sufficient condition for* (4.1) *to admit an isotropic solution is* (4.4) *and* $R = 16\pi(\rho_{00} + \Lambda/4\pi)$ *where* α, $U^h U_h$ *and* p_0 *are determined by* (4.2), (4.5) *and* (4.7).

We have sketched the proof for the necessary condition of the theorem. The sufficient condition can be established as in the case for the field in empty space.

The general results of the paper were obtained when the author was in Peiping before the Summer of 1936 and the details have been completed here in Princeton.

It is a pleasure for him to express his gratitude to National Tsing Hua University for granting the sabbatical leave 1936-1937 and to the Institute for Advanced Study and Princeton University for their hospitality and the facilities of research put at his disposal.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

ON TENSORS RELATIVE TO THE EXTENDED POINT TRANSFORMATION.*

By H. V. Craig.

1. Introduction. The transformation equations of the quantities dx^r/dt and $\partial f/\partial x^r$ may be said to lead to the tensor concept. Quite similarly the generalizations of these vectors, namely, $x^{(a+1)r}$, $\partial f(x \cdot x', \cdots, x^{(M)})/\partial x^{(a)r}$, $(x^{(a)r} = d^a x^r/dt^a)$ suggest a development which at first sight appears to be an extension of this very important notion. More properly perhaps the new quantities "extensors" should be regarded merely as special tensors—tensors relative to a rather special transformation. However, because of their special character a process occurs (contraction over a reduced range) which does not appear in ordinary tensor analysis.

No problem is involved in incorporating the developments of ordinary tensor analysis and consequently our interest lies elsewhere. Two subjects which are perhaps worthy of investigation are: (a) the systemization of certain of the results appearing in the various higher order geometries,[1] for example, we shall show that most of the invariants given may be regarded as formed by contraction over a reduced range; (b) the construction of tensors from extensors.

2. Notation. Following Schouten we shall use but one root letter x for all coördinate systems and distinguish between different coördinate systems by means of the letters employed as indices. We shall have occasion to consider simultaneously at most three different coördinate systems, let us call them temporarily x, y, and z, and shall adopt the associations: $x: r, s, t$; $y: i, j, k$; $z: u, v, w$. Thus. $x^r = x^r$; $x^i = y^i$; $x^u = z^u$. Likewise if V is a vector then V^r will stand for the r-th component of V in the x coördinate system, while V^i will denote the i-th component of V in the y system etc. Similarly, with regard to tensors and extensors, the presence as an index of one or more of

* Received March 10, 1937; Revised (notation only) July 13, 1937.

[1] I. e. geometries whose metrics depend on derivatives of higher order than the first. A bibliography of papers in this field is appended. The geometry studied by J. L. Synge is apparently the most general being based on a metric $F(t, x, x', \cdots, x^{(M)})$. A. Kawaguchi and others have investigated metrics which do not depend on t explicitly, whereas, the present writer has confined his attention primarily to the very special case in which Fdt is invariant under parameter change.

the letters i, j, k will serve to indicate that the component involved belongs to the y coördinate system etc.

Derivatives with respect to the parameter t will be designated for the most part by means of enclosed indices, whereas partial derivatives will be indicated by subscripts, specifically:

$$
\begin{aligned}
&x'^r = dx^r/dt; \quad x^{(\alpha)r} = d^\alpha x^r/dt^\alpha; \quad F_{(\alpha)r} = \partial F/\partial x^{(\alpha)r};\\
&X_i{}^r = X^{(0)r}_{(0)i} = \partial x^r/\partial x^i; \quad X^{(\alpha)r}_{(\gamma)i} = \partial x^{(\alpha)r}/\partial x^{(\gamma)i};\\
&X_i{}^{r(\alpha)} = d^\alpha X_i{}^r/dt^\alpha;
\end{aligned}
\tag{2.1}
$$

$\binom{u}{v}$ is a binomial coefficient.

In addition to the dimensionality of the space, N, we shall require another invariant M—the order of the space. The summations indicated by repeated lower case Latin indices shall be from one to N, whereas those represented by a repetition of small Greek indices shall be from zero to M unless otherwise specified. Whenever the range of summation of a Greek index is not zero to M a summation sign will be introduced and the range indicated in the usual way. It should be borne in mind that quantities such as $X^{(\alpha)r}_{(\beta)i}$ have the value zero if $\alpha < \beta$ and consequently there is some latitude in the range of summations involving these derivatives. Incidentally, in all but a very few instances the irregularities of range could be removed by defining all symbols bearing negative indices to have the value zero.

Following McConnell we shall use *capital* letters for repeated indices which are *not* to be summed.

Finally, the various scalars, vectors, and extensors which are to appear will be assumed to be absolute rather than relative and to be functions of the coördinate variables and their derivatives with respect to the parameter t.

3. Extensors. The point transformation: $x^r = x^r(y^1, y^2, \cdots, y^N)$, which we assume to be of class ω and regular,[2] gives rise upon successive differentiation to the "extended point transformation":

$$
\begin{cases}
x^r = x^r(y); \quad x'^r = X_i{}^r x'^i;\\
x''^r = X_i{}^r x''^i + X_{ij}{}^r x'^i x'^j; \quad x^{(M)r} = X_i{}^r x^{(M)i} + \cdots;
\end{cases}
\tag{3.1}
$$

and the notion of extensor is merely that of tensor relative to this extended transformation. As particular examples of these quantities and as an illustration of the notation adopted let us consider the transformation equations of $x^{(\alpha+1)r}$ or $x'^{(\alpha)r}$, and $F_{(\alpha)r}$.

---

[2] See Veblen and Whitehead, *The Foundations of Differential Geometry*, London, Cambridge, chap. 3.

Examining (3.1) we note that $x^{(\alpha)r}$ may be expressed in terms of $x^{(\beta)i}$, $\beta = 0, \cdots, \alpha$, hence

$$(x^{(\alpha)r})' = X^{(\alpha)r}_{(\beta)i}(x^{(\beta)i})' \tag{3.2}$$

or otherwise expressed

$$x'^{r(\alpha)} = x'^{i(\beta)}X^{(\alpha)r}_{(\beta)i} = \sum_{i=1}^{N}\sum_{\beta=0}^{M} x'^{i(\beta)}X^{(\alpha)r}_{(\beta)i}; \quad (X^{(\alpha)r}_{(\beta)i} = 0, \text{ for } \beta > \alpha). \tag{3.3}$$

This result may be generalized by means of the very useful formula [3]

$$X^{(A)r}_{(B)i} = \binom{A}{B} X_i{}^{r(A-B)} \tag{3.4}$$

thus if V is a sufficiently differentiable vector

$$V^{r(A)} = (X_i{}^r V^i)^{(A)} = \sum_{\beta=0}^{A}\binom{A}{\beta} X_i{}^{r(A-\beta)}V^{i(\beta)} = X^{(A)r}_{(\beta)i}V^{i(\beta)}. \tag{3.5}$$

Finally, the transformation equation for $F_{(\alpha)r}$, F being a differentiable scalar, is

$$F_{(\beta)i} = F_{(\alpha)r}X^{(\alpha)r}_{(\beta)i}, \tag{3.6}$$

and we are now ready to turn to the definition of extensor confining our statement for the sake of brevity to a third order quantity.

Definition. If there is given at a point of a curve of class C^M a set of $n^3(M+1)^2$ labelled numbers or components in each coördinate system of our group and if further the components $T^{ar}_{s.\beta t}$ associated with any coördinate system x^r are related to the components $T^{\gamma i}_{j.\delta k}$ belonging to any other system x^i according to the transformation equation

$$T^{ar}_{s.\beta t} = T^{\gamma i}_{j.\delta k}X^{(\alpha)r}_{(\gamma)i}X_s{}^j X^{(\delta)k}_{(\beta)t} \tag{3.7}$$

then we shall speak of these labelled numbers as the components of a mixed third order extensor—excontravariant of order one, covariant of order one, and excovariant of order one.

In particular if $V_{ar} = F_{(a)r}$, we shall speak of V_{ar} as the exgradient of F; likewise the quantities $x'^{r(a)}$ might be regarded as the components of a tangent vector in the extended sense.

The questions of consistency which arise in connection with (3.7) may

[3] See H. V. Craig, 5, p. 457. Capital indices it will be recalled are not summed.

be disposed of as in ordinary tensor analysis. Obviously, the extended transformation (3.1) admits of a unique inverse and further examination reveals that the Jacobian is $(\partial x/\partial y)^{M+1}$, hence this transformation is regular. Furthermore, we may write

$$X^{(\alpha)r}_{(\beta)i}X^{(\beta)i}_{(\gamma)s} = \delta_\gamma{}^\alpha \delta_s{}^r \tag{3.8}$$

and

$$X^{(\alpha)r}_{(\gamma)u} = X^{(\alpha)r}_{(\beta)i}X^{(\beta)i}_{(\gamma)u}. \tag{3.9}$$

These formulas which are evident from the $(M+1)N$ dimensional viewpoint, can also be obtained by applying [4] (3.4) to the identity

$$X_u{}^{r(A-\Gamma)} = (X_i{}^r X_u{}^i)^{(A-\Gamma)} = \sum_{\beta=0}^{A-\Gamma} \binom{A-\Gamma}{\beta} X_i{}^{r(A-\Gamma-\beta)} X_u{}^{i(\beta)}. \tag{3.10}$$

Obviously, the rules relating to the sum and product of tensors carry over. Likewise, the contraction of an extensor yields one of lower order thus

$$T^{\alpha r}_{s.\alpha r} = T^{\gamma i}_{j.\delta k} X^{(\alpha)r}_{(\gamma)i} X_s{}^j X^{(\delta)k}_{(\alpha)r} = T^{\gamma i}_{j.\gamma i} X_s{}^j. \tag{3.11}$$

A very simple example of contraction is furnished by the formula

$$F' = F_{(\alpha)r} x'^{(\alpha)r}. \tag{3.12}$$

More, generally, if F is any differentiable scalar and V any differentiable vector then $F_{(\alpha)r}$ and $V^{r(\alpha)}$ are extensors and consequently $F_{(\alpha)r}V^{r(\alpha)}$ is an invariant. However, the quantities $F_{(\alpha)i(\beta)j}$ are not the components of an extensor, since they follow the transformation rule

$$F_{(\alpha)i(\beta)j} = F_{(\gamma)r(\epsilon)s} X^{(\gamma)r}_{(\alpha)i} X^{(\epsilon)s}_{(\beta)j} + F_{(\gamma)r} X^{(\gamma)r}_{(\alpha)i(\beta)j}. \tag{3.13}$$

Incidentally, $X^{(\Gamma)r}_{(A)i(\beta)j}$ may be written in the form $\binom{\Gamma}{A} X_i{}^{r(\Gamma-A)}_{(\beta)j}$ and consequently it vanishes whenever β exceeds $\Gamma - A$.

The matter of eliminating the second derivatives from (3.13) may be treated just as in tensor analysis. Thus, let $g_{\alpha i\beta j}$ be a second order extensor of non-vanishing determinant. The first N terms in the first row of this determinant would be the quantities $g_{01.0j}$, while the second group of N would consist of $g_{01.1j}$ etc. By forming the derivatives $g_{\alpha i.\beta j.(\gamma)k}$, Christoffel symbols and an affine connection could be constructed in the usual way. Hence, out of a scalar or extensor, extensors of any order may be constructed by covariant differentiation and finally invariants by contraction.

---

[4] First replace β with $\beta - \Gamma$ and note that the new limits of summation may be taken to be $\beta = 0$, $\beta = M$.

4. Extensors of reduced range. Examining equation (3.7) and recalling that $X^{(\delta)k}_{(\beta)t}$ is equal to zero whenever β exceeds δ, we observe that the components of $T^{\alpha r}_{s.\beta t}$ which have β equal to or greater than b depend only on the similarly labelled components $T^{\gamma i}_{j.\delta k}$ i. e. on those for which δ is not less than b. The collection of all such components will be referred to as the components of an extensor of reduced range. For example, the covariant tensor $F_{(M)r}$ is an extensor of reduced range ($b = M$) and likewise the quantities $F_{(M-1)r}$, $F_{(M)r}$ ($b = M-1$). Similarly, the range of a superscript may be reduced. Thus, returning to (3.7) we note that the $T^{\alpha r}_{s.\beta t}$ having α equal to or less than a depend only on the corresponding components $T^{\gamma i}_{j.\delta k}$, $\gamma \leq a$. A particular example is furnished by the derivatives of a contravariant vector V, thus

$$V^{i(\beta-E)} = \sum_{\gamma=E}^{M} V^{r(\gamma-E)} X^{(\beta-E)i}_{(\gamma-E)r}; \qquad M \geq \beta \geq E. \tag{4.1}$$

If E is equal to M, the range is completely reduced and (4.1) becomes a tensor transformation.

Although, differentiation of equation (3.7) yields second order derivatives we can employ this process over a restricted range to construct reduced extensors. As an illustration let us consider the following theorem which obviously admits of generalization.

THEOREM (4.1). *If $T^{\alpha i}_{\beta j}$ are the components of an extensor of reduced range ($\alpha \leq a$, $\beta \geq b$) then for $\gamma > a$ and $\gamma > M - b$ simultaneously, the set of quantities $T^{\alpha i}_{\beta j.(\gamma)k}$ constitute the components of an extensor of reduced range.*

Proof. The law of transformation of $T^{\alpha i}_{\beta j}$ is expressed by the relationship;

$$T^{\alpha i}_{\beta j} = T^{\sigma r}_{\delta s} X^{(\alpha)i}_{(\sigma)r} X^{(\delta)s}_{(\beta)j}, \qquad \alpha \leq a,\ \beta \geq b. \tag{4.2}$$

Differentiating (4.2) with respect to $x^{(\gamma)k}$, $\gamma > a$, $\gamma > M - b$, we obtain the desired equality

$$T^{\alpha i}_{\beta j.(\gamma)k} = T^{\sigma r}_{\delta s.(\epsilon)t} X^{(\epsilon)t}_{(\gamma)k} X^{(\alpha)i}_{(\sigma)r} X^{(\delta)s}_{(\beta)j}, \quad \alpha \leq a,\ \beta \geq b,\ \gamma > a,\ \gamma > M - b, \tag{4.3}$$

at once, since $X^{(\alpha)i}_{(\sigma)r.(\epsilon)t} X^{(\epsilon)t}_{(\gamma)k}$ and $X^{(\delta)s}_{(\beta)j.(\gamma)k}$ vanish by virtue of the restrictions on α, β, and γ.

An illustration of this process is furnished by the differentiation of a tensor. For example, the quantities $T^{i}_{jk.(\beta)l}$, $\beta > 0$, are the components of a reduced extensor.

Finally, with regard to the differentiation of (4.2) it should be noted that if γ were permitted to take on the larger of the two values a and $M - b$ or

their common value if $a = M - b$ then there would appear in the right member of (4.3) one or more terms involving mixed partial derivatives of the unprimed variables. These derivatives could be eliminated by means of a zero-th order affine connection—the resulting quantities are however irregular and of no apparent interest.

5. Contraction over a reduced range. The existence of extensors of reduced range suggests the possibility of constructing invariants and tensors by contraction over a reduced range. As a first trial it would be natural to consider the expression $\sum_{\beta=E}^{M} F_{\beta i} V^{\beta-E,i}$. Investigation however reveals that a factor $\binom{\beta}{E}$ is needed. The general process is exemplified by the theorem which follows.

THEOREM (5.1). *If $T_{k.}^{ai.\beta-E,j}_{\gamma l}$ is an extensor of reduced range: $E \leq \beta \leq M$, $E \leq \gamma \leq M$, then the quantities $\sum_{\beta=E}^{M} \binom{\beta}{E} T_{k.}^{ai.\beta-E,j}_{\beta j}$ are the components of a mixed extensor—covariant of order one and excontravariant of order one.*

Proof. Evidently the transformation equation to be investigated is

$$\sum_{\beta=E}^{M} \binom{\beta}{E} T_{k.}^{ai.\beta-E,j}_{\beta j} = \sum_{\beta=E}^{M}\sum_{\delta=E}^{M}\sum_{\rho=E}^{M} T_{t.}^{\gamma r.\delta-E,s}_{\rho q} X_{(\alpha)r}^{(\alpha)i} X_k{}^t \binom{\beta}{E} X_{(\delta-E)s}^{(\beta-E)j} X_{(\beta)j}^{(\rho)q}. \tag{5.1}$$

Now by virtue of (3.4) and the equality of $\binom{B}{E}\binom{B-E}{\Delta-E}\binom{B}{\Delta}^{-1}$ and $\binom{\Delta}{E}$, we may write successively,

$$\binom{B}{E} X_{(\Delta-E)s}^{(B-E)j} = \binom{B}{E}\binom{B-E}{\Delta-E} X_s{}^{j(\beta-\Delta)} = \binom{B}{E}\binom{B-E}{\Delta-E}\binom{B}{\Delta}^{-1} X_{(\Delta)s}^{(B)j} = \binom{\Delta}{E} X_{(\Delta)s}^{(B)j} \tag{5.2}$$

and consequently

$$\sum_{\beta=E}^{M} \binom{\beta}{E} X_{(\Delta-E)s}^{(\beta-E)j} X_{(\beta)j}^{(\rho)q} = \sum_{\beta=E}^{M} \binom{\Delta}{E} X_{(\Delta)s}^{(\beta)j} X_{(\beta)j}^{(\rho)q}. \tag{5.3}$$

For $\Delta \geq E$ the range of β may be taken to be 0 to M and hence the last summation sign may be dropped. Thus the last member of (5.3) reduces to $\binom{\Delta}{E}\delta_\Delta{}^\rho \delta_s{}^q$ and the right member of (5.1) becomes

$$\sum_{\delta=E}^{M} \binom{\delta}{E} T_{t.}^{\gamma r.\delta-E,s}_{\delta s} X_{(\alpha)r}^{(\alpha)i} X_k{}^t \tag{5.4}$$

and the theorem is established.

The quotient law for contraction over a reduced range is illustrated by the next theorem.

THEOREM (5.2). *If for each differentiable vector V the set of quantities* $\sum_{\beta=E}^{M}\binom{\beta}{E}T^{j}{}_{\alpha i\,.\,\beta k}V^{k(\beta-E)}$ *are the components of a mixed extensor of the type indicated by the free indices then* $T^{j}{}_{\alpha i\,.\,\beta k}$ *is a third order mixed extensor of the reduced range* $E \leq \beta \leq M$.

Proof. By virtue of the hypotheses we have

$$(5.5)\qquad \begin{aligned}&\sum_{\beta=E}^{M}\binom{\beta}{E}T^{j}{}_{\alpha i\,.\,\beta k}V^{k(\beta-E)}\\ &=\sum_{\delta=E}^{M}\binom{\delta}{E}T^{s}{}_{\gamma r\,.\,\delta t}V^{t(\delta-E)}X_{s}{}^{j}X^{(\gamma)r}_{(\alpha)i}\\ &=\sum_{\beta=E}^{M}\sum_{\delta=E}^{M}\binom{\delta}{E}T^{s}{}_{\gamma r\,.\,\delta t}V^{k(\beta-E)}X^{(\delta-E)t}_{(\beta-E)k}X_{s}{}^{j}X^{(\gamma)r}_{(\alpha)i}\end{aligned}$$

and consequently

$$(5.6)\qquad \sum_{\beta=E}^{M}\left[\binom{\beta}{E}T^{j}{}_{\alpha i\,.\,\beta k}-\sum_{\delta=E}^{M}\binom{\delta}{E}T^{s}{}_{\gamma r\,.\,\delta t}X^{(\delta-E)t}_{(\beta-E)k}X_{s}{}^{j}X^{(\gamma)r}_{(\alpha)i}\right]V^{k(\beta-E)}\equiv 0.$$

By substituting for V^k successively the quantities $\delta_l{}^k$, $\delta_l{}^k t$ etc., t being the parameter, we learn that the bracket in (5.6) must vanish for each admissible value of k and β. Furthermore, by way of (5.2) we may replace $\binom{\Delta}{E}X^{(\Delta-E)t}_{(B-E)k}$ by $\binom{B}{E}X^{(\Delta)t}_{(B)k}$ and the theorem follows.

6. The structure of certain tensors. The most important scalar equations of generalized geometry are due to Zermelo and express the various conditions for the invariance of Fdt, $F=F(x,x',\cdots,x^{(M)})$, under change of parameter. It may be of interest to examine their structure in the light of the foregoing developments. The equations[5] are:

$$(6.1)\qquad (x'^{r}T)^{(a)}F_{(a)r}=(FT)';$$

$$(6.2)\qquad \sum_{a=\Gamma}^{M}\binom{a}{\Gamma}x'^{r(a-\Gamma)}F_{(a)r}=\delta_{\Gamma}{}^{1}F;$$

$$(6.3)\qquad \sum_{a=1}^{M}\binom{a}{1}x'^{r(a-1)}Z_{ar}=F;$$

$$(6.4)\qquad (Z_{Ar}=1/A\cdot\sum_{\beta=A}^{M}(-1)^{\beta-A}F_{(\beta)r}{}^{(\beta-A)},\qquad (A=1,\cdots,M)).$$

[5] See A. Kneser, *Lehrbuch der variationsrechnung* (1900), p. 195. Also, H. V. Craig, 1, p. 559; 5, p. 461; 6, p. 835.

The left members of the first two of the foregoing set are obviously examples of complete contraction and contraction over a reduced range. The third equation suggests the following theorem.

THEOREM (6.1). *The set of quantities* $Z_{\alpha r}$, $1 \leq \alpha \leq M$, *constitute the components of an extensor of reduced range.*

Proof. One verifies readily that for the ranges: $\alpha = M$; $M-1 \leq \alpha \leq M$; the quantities $Z_{\alpha r}$ are indeed the components of a reduced extensor and we proceed to establish the theorem by induction. Thus, we shall assume the theorem for the range $A+1 \leq \alpha \leq M$ and on this basis deduce that Z_{Ar} obeys the desired transformation equation. Now it may be shown that these quantities obey the recursion formula [6]

$$Z_{Ai} = 1/A \cdot F_{(A)i} - (A+1)/A \cdot Z_{A+1,i}' \tag{6.5}$$

and consequently it will suffice to investigate the transformation of the right member of (6.5).

By virtue of the extensor character of $F_{(A)i}$ and $Z_{A+1,i}$ we may write

$$F_{(A)i} = \sum_{\gamma=A}^{M} F_{(\gamma)r} X^{(\gamma)r}_{(A)i}, \tag{6.6}$$

$$Z_{A+1,i} = \sum_{\gamma=A+1}^{M} Z_{\gamma r} X^{(\gamma)r}_{(A+1)i} \tag{6.7}$$

and, differentiating (6.7) with respect to the parameter t,

$$Z_{A+1,i}' = \sum_{\gamma=A+1}^{M} Z_{\gamma r}' X^{(\gamma)r}_{(A+1)i} + \sum_{\gamma=A+1}^{M} Z_{\gamma r} X^{(\gamma)r}_{(A+1)i}{}'. \tag{6.8}$$

If we replace γ with $\gamma+1$ in the first term of the right member of (6.8) we get

$$Z_{A+1,i}' = \sum_{\gamma=A}^{M-1} Z_{\gamma+1,r}' X^{(\gamma+1)r}_{(A+1)i} + \sum_{\gamma=A+1}^{M} Z_{\gamma r} X^{(\gamma)r}_{(A+1)i}{}'. \tag{6.9}$$

Furthermore, by way of (6.5), (6.6), and (6.9) we have

$$\begin{aligned} Z_{Ai} = 1/A \sum_{\gamma=A}^{M} F_{(\gamma)r} X^{(\gamma)r}_{(A)i} - (A+1)/A \sum_{\gamma=A}^{M-1} Z_{\gamma+1,r}' X^{(\gamma+1)r}_{(A+1)i} \\ - (A+1)/A \sum_{\gamma=A+1}^{M} Z_{\gamma r} X^{(\gamma)r}_{(A+1)i}{}'. \end{aligned} \tag{6.10}$$

Evidently, the lower limit of the last summation can be changed to A while

[6] See H. V. Craig, 1, p. 560.

$-(A+1)/A \cdot X^{(\gamma+1)r}_{(A+1)i}$ can be replaced by $-(\gamma+1)/A \cdot X^{(\gamma)r}_{(A)i}$, $-(A+1)/A \cdot X^{(\gamma)r}_{(A+1)i}{}'$ by $(1-\gamma/A)X^{(\gamma)r}_{(A)i}$, and Z_{Mr} by $1/M \cdot F_{(M)r}$.

This done the right member of (6.10) assumes the form

$$(6.11)\quad \sum_{\gamma=A}^{M-1}\{\gamma/A[1/\gamma \cdot F_{(\gamma)r} - (\gamma+1)/\gamma \cdot Z_{\gamma+1,r}{}'] + (1-\gamma/A)Z_{\gamma r}\}X^{(\gamma)r}_{(A)i} + (1/A \cdot F_{(M)r} + (1-M/A)1/M \cdot F_{(M)r})X^{(M)r}_{(A)i}$$

and thus by way of (6.5) equation (6.10) reduces to

$$(6.12)\quad Z_{Ai} = \sum_{\gamma=A}^{M-1} Z_{\gamma r}X^{(\gamma)r}_{(A)i} + Z_{Mr}X^{(M)r}_{(A)i}$$

and our theorem is proved and the structure of (6.3) established.

As a final scalar let us consider the Synge invariant;[7]

$$(6.13)\quad \sum_{\alpha=0}^{M-1} V^{r(\alpha)} \sum_{\beta=1}^{M-\alpha} (-1)^{\beta-1} F^{\ldots(\beta-1)}_{\ldots(\alpha+\beta)r}.$$

If we replace first α with $\alpha-1$ and then β by $\beta-\alpha+1$ the expression (6.13) may be written in the form

$$(6.14)\quad \sum_{\alpha=1}^{M}\binom{\alpha}{1} V^{r(\alpha-1)} \cdot 1/\alpha \cdot \sum_{\beta=\alpha}^{M} (-1)^{\beta-\alpha} F^{\ldots(\beta-\alpha)}_{\ldots(\beta)r}$$

which by virtue of the definition of Z_{ar} reduces to the contraction

$$(6.15)\quad \sum_{\alpha=1}^{M}\binom{\alpha}{1} V^{r(\alpha-1)}Z_{\alpha r}.$$

Among the tensors whose character is not apparent from the N-dimensional viewpoint are the quantities $\overset{p-R}{D}_{rs}(T)V^s$, due to Kawaguchi,[8]

$$(6.16)\quad \overset{p-R}{D}_{rs}(T)V^s = \sum_{\alpha=R}^{p}\binom{\alpha}{R} T_{r(\alpha)s}V^{s(\alpha-R)}.$$

This is obviously another example of contraction over a reduced range.

Among the most interesting quantities of higher order geometry are the vectors $\overset{A}{E}_r$[9]

$$(6.17)\quad \overset{A}{E}_r = \sum_{\gamma=\alpha}^{M} (-1)^{\gamma}\binom{\gamma}{A} F^{\ldots(\gamma-A)}_{\ldots(\gamma)r}.$$

Examination of their structure suggests that these vectors may be generalized

[7] See J. L. Synge, 1, p. 683.

[8] See A. Kawaguchi, 4, p. 149.

[9] See J. L. Synge, 1, p. 684; H. V. Craig, 6, p. 833.

by replacing the exgradient $F_{(\gamma)r}$ with an arbitrary differentiable extensor $V_{\gamma r}$. As an aid to this investigation we introduce the identity

$$(6.18)\qquad \sum_{\mu=A}^{\Gamma}\binom{\mu}{A}(-1)^{\mu}(FX^{(\Gamma)r}_{(\mu)i})^{(\mu-A)}=\binom{\Gamma}{A}(-1)^{\Gamma}F^{(\Gamma-A)}X_i{}^r$$

which may be established as follows.

By virtue of the rule for differentiating a product and the formula (3.4) we may express the left member of (6.18) in the form

$$(6.19)\qquad \sum_{\mu=A}^{\Gamma}\binom{\mu}{A}(-1)^{\mu}\sum_{\sigma=0}^{\mu-A}\binom{\mu-A}{\sigma}F^{(\sigma)}\binom{\Gamma}{\mu}X_i{}^{r(\Gamma-A-\sigma)}.$$

Interchanging the summation signs and replacing $\binom{M}{A}\binom{M-A}{\Sigma}\binom{\Gamma}{M}$ by the equivalent product $\binom{\Gamma}{A}\binom{\Gamma-A}{\Sigma}\binom{\Gamma-A-\Sigma}{\Gamma-M}$, we obtain the expression,

$$(6.20)\qquad \sum_{\sigma=0}^{\Gamma-A}\binom{\Gamma}{A}\binom{\Gamma-A}{\sigma}F^{(\sigma)}X_i{}^{r(\Gamma-A-\sigma)}\sum_{\mu=\sigma+A}^{\Gamma}(-1)^{\mu}\binom{\Gamma-A-\sigma}{\Gamma-\mu}.$$

Now, if we replace μ with $\mu+\sigma+A$ and denote $\Gamma-A-\sigma$ by b, the last summation may be written in the form $(-1)^{\sigma+A}\sum_{\mu=0}^{b}(-1)^{\mu}\binom{b}{\mu}$ and consequently reduces to $(-1)^{\sigma+A}\delta_0{}^b$. Hence (6.20) reduces to $(-1)^{\Gamma}\binom{\Gamma}{A}F^{(\Gamma-A)}X_i{}^r$, the right member of (6.18). This accomplished we are ready to consider the theorem:

THEOREM (6.2). *If $V_{\gamma r}$ is a differentiable extensor then the quantities*

$$(6.21)\qquad \sum_{\gamma=A}^{M}(-1)^{\gamma}\binom{\gamma}{A}V^{(\gamma-A)}_{\gamma r}$$

are the components of a vector.

Proof. If in (6.18) we replace Γ with γ, F with $V_{\gamma r}$ and sum both members from $\gamma=A$ to $\gamma=M$, we obtain (after changing the order of summation on the left)

$$(6.22)\qquad \sum_{\mu=A}^{M}(-1)^{\mu}\binom{\mu}{A}\left(\sum_{\gamma=\mu}^{M}V_{\gamma r}X^{(\gamma)r}_{(\mu)i}\right)^{(\mu-A)}=\sum_{\gamma=A}^{M}(-1)^{\gamma}\binom{\gamma}{A}V^{(\gamma-A)}_{\gamma r}X_i{}^r$$

or

$$(6.23)\qquad \sum_{\mu=A}^{M}(-1)^{\mu}\binom{\mu}{A}V^{(\mu-A)}_{\mu i}=\sum_{\gamma=A}^{M}(-1)^{\gamma}\binom{\gamma}{A}V^{(\gamma-A)}_{\gamma r}X_i{}^r,$$

and our theorem is established.

TEXAS UNIVERSITY.

BIBLIOGRAPHY.

E. Cartan, "La géométrie de l'integral $\int F(x, y, y', y'')\,dx$," *Journal de mathématiques pures et appliques*, vol. 15 (1936), p. 42.

H. V. Craig, 1, "On the solution of the Euler equations for their highest derivatives," *Bulletin of the American Mathematical Society*, vol. 36 (1930), p. 560; 2, "On parallel displacement in a non-Finsler space," *Transactions of the American Mathematical Society*, vol. 33 (1931), p. 129; 3, "On a covariant differentiation process," *Bulletin of the American Mathematical Society*, vol. 37 (1931), p. 731; 4, "On a covariant differentiation process, II," *Bulletin of the American Mathematical Society*, vol. 39 (1933), p. 919; 5, "On a generalized tangent vector," *The American Journal of Mathematics*, vol. 57 (1935), p. 457; 6, "On a generalized tangent vector, II," *The American Journal of Mathematics*, vol. 58 (1936), p. 833.

S. Hokari, "Die Geometrie des Integrals $\int (a_i a_j x''^i x''^j + 2b a_i x''^i + c)^{1/p} dt$," *Proceedings of the Imperial Academy of Tokio*, vol. 12 (1936), p. 209.

H. Hombu, 1, "On a non-Finsler metric space," *The Tôhoku Mathematical Journal*, vol. 37 (1933), p. 190; 2, "Invariantentheorie des Integrals $\int F(x, y, y', y'', y''')\,dx$," *Proceedings of the Imperial Academy of Tokio*, vol. 12 (1936), p. 156.

A. Kawaguchi, 1, "Theory of connections in the generalized Finsler Manifold," *Proceedings of the Imperial Academy of Tokio*, vol. 7 (1931), p. 211; 2, "Die Differentialgeometrie in der varallgemeinerten Mannigfaltigkeit," *Rendiconti Circolo Matematico di Palermo*, vol. 56 (1932), p. 245; 3, "Theory of connections in the generalized Finsler Manifold, III," *Proceedings of the Imperial Academy of Tokio*, vol. 9 (1933), p. 347; 4, "Some intrinsic derivations in a generalized space," *Proceedings of the Imperial Academy of Tokio*, vol. 12 (1936), p. 149; 5, "Certain identities in a generalized space," *Proceedings of the Imperial Academy of Tokio*, vol. 12 (1936), p. 152; 6, "Die Geometrie des Integrals $\int (A_i x''^i + B)^{1/p} dt$," *Proceedings of the Imperial Academy of Tokio*, vol. 12 (1936), p. 205.

D. D. Kosambi, "Path-spaces of higher order," *The Quarterly Journal of Mathematics*, Oxford Series, vol. 7 (1936), p. 97.

J. L. Synge, "Some intrinsic and derived vectors in a Kawaguchi space," *The American Journal of Mathematics*, vol. 57 (1935), p. 679.

A SERIES OF INVOLUTORIAL CREMONA TRANSFORMATIONS IN S_n BELONGING MULTIPLY TO A NON-LINEAR LINE COMPLEX.*

By VIRGIL SNYDER and EVELYN CARROLL-RUSK.

The involutorial transformation I will first be developed in three way projective space S_3, then extended to S_n.

1. Definition of the involution. Consider a pencil

$$\mu_2 H_1(x) - \mu_1 H_2(x) = 0 \tag{1}$$

of quadric surfaces, with a proper or composite curve C_4 for base, and a rational ruled surface R_{n+m} of order $n+m$, generated by the line

$$(ax) = 0, \qquad (bx) = 0, \tag{2}$$

in which each coefficient a_i is a binary form in λ_1, λ_2 of order m_1, and each coefficient b_i is a binary form in λ_1, λ_2 of order m_2.

Let ϕ_1, ϕ_2 be two binary forms in μ_1, μ_2 of degree k, and ψ_1, ψ_2 be two quadratic forms in λ_1, λ_2.

Between the ϕ, ψ is the relation

$$\phi_1/\phi_2 = \psi_1/\psi_2. \tag{3}$$

With these premises the involution I can be defined as follows: Given any point (y). By (1) it determines a quadric of the pencil, hence $\mu_1 = H_1(y)$, $\mu_2 = H_2(y)$.

These values substituted in (3) define λ_1/λ_2 as roots of a quadratic equation and consequently, by (2), a pair of generators of R_{n+m} belonging to a g_2^1.

Through y can be drawn one transversal meeting both, and it will meet the quadric $H(y)$ determined by (y) in a second point (y'), conjugate of (y) in I.

2. Properties of the involution. Fundamental elements. From (3) the quadrics of the pencil (1) are arranged in sets of k, all associated with the same pair of generators of $R_{m_1+m_2}$. Every quadric of the pencil is invariant

* Presented to the American Mathematical Society at New York, March 27, 1937. The first part was also read before the International Congress of Mathematicians, Oslo, July 13, 1936. Received by the Editors May 3, 1937.

under the involution I, and every line joining a pair PP' of conjugate points contains k such pairs. For each set of k quadrics the transformation is expressed by means of a linear line congruence.

The fundamental elements are of three kinds:

(a) the basic quartic C_4 of the pencil of quadrics.

From a point P on C_4 can be drawn a transversal t to every pair of conjugate generators of $R_{m_1+m_2}$. Each such transversal determines a set of k quadrics, and each quadric has one point on the transversal apart from P. The locus of this point as t describes a cone with vertex P is the image of P. As P describes C_4, the curve generates a surface K, the principal element, image of the curve C_4.

(b) Let g_1, g_2 be the generators of $R_{m_1+m_2}$ associated with the quadrics $H_1, \cdots, H_k$. Each line meets all k associated quadrics each in two points. Let Q be a point g_1, H_i. The plane Q, g_2 meets H_i in a conic passing through Q; this conic is the image of Q. As g describes $R_{m_1+m_2}$ Q describes a curve γ which is double on the web of surfaces conjugate to the planes of space, and it lies on the surface of invariant points. As Q describes γ, its conjugate conic describes the principal surface Γ.

These are the only fundamental curves of the first kind. The surfaces K, Γ comprise the Jacobian of the web.

The sets of k quadrics associated with coincident lines $g_1 = g_2$ contain an infinite number of parasitic conics, hence appear as factors and do not affect the proper transformation.

(c) The common bisecants of γ and C_4 are parasitic. Each point of every such line is transformed into the whole line. All these lines lie on the surface σ of invariant points.

(d) For certain values of μ the associated lines g_1, g_2 intersect. Their plane meets the associated k quadrics in k conics of a pencil, each of which is parasitic. These conics all lie on the surface σ of invariant points.

The lines (c) and the conics (d) include all the fundamental curves of the second kind.

3. Analytic procedure. Let the line t meet g_1 and z' and g_2 in z''. By (3) the coördinates of $z = z' + z''$ can be expressed rationally in terms of μ. Each z_i is of order $2p+1$ in (y) and contains μ to multiplicity p, where $p = k(m_1 + m_2 - 1)$. Then $y'_i = \sigma z_i + \tau y_i$, wherein

$$\sigma = 2(H_1(y)H_2(y,z) - H_2(y)H_1(y,z)),$$
$$\tau = -(H_1(y)H_2(z) - H_2(y)H_1(z)),$$
$$H_i(y,z) = H_i(z,y) \text{ is the polar of } H(y) \text{ as to } z.$$

The surface $\sigma = 0$ is the locus of points invariant under I. Each point of σ is a point of contact of a line of the linear congruence g_1, g_2 which touches an associated quadric.

The equations of the curve γ can be found by eliminating λ, λ' between (2) and (3). From (1), (3) may be written in the form

$$\{A_1\phi_1(H_1, H_2) + B_1\phi_2(H_1, H_2)\}\lambda_1^2 + \{A_2\phi_1(H_1, H_2) + B_2\phi_2(H_1, H_2)\}\lambda_1\lambda_2 + \cdots = 0.$$

Eliminate λ between this and $(ax) = 0$. The result is a surface $F_{2km_1+2} : C_4^{km_1}$. Similarly, by eliminating λ between the same quadratic equation and $(bx) = 0$, we obtain $F_{2km_2+2} : C_4^{km_2}$. Each of these surfaces contains γ and C_4;

$$(2km_1 + 2)(2km_2 + 2) - 4k^2m_1m_2 = 4[k(m_1 + m_2) + 1],$$

but each line of R has two points on γ, hence the order of γ is $2p + 2 + 2k$. It lies on the surface $R_{m_1+m_2}$.

Let $(ax)(by) - (ay)(bx) = \Sigma P_i x_i = 0$, and $\Sigma P'_i x_i = 0$ be the equations of t, wherein $P_i = \Sigma p_{ik} y_k$, $p_{ik} = a_i b_k - a_k b_i$. This line meets $\Sigma x_i = 0$ in the point v, each coördinate v_i being obtained from the matrix

$$\begin{Vmatrix} P_1 & P_2 & P_3 & P_4 \\ P'_1 & P'_2 & P'_3 & P'_4 \\ 1 & 1 & 1 & 1 \end{Vmatrix}$$

by suppressing the i-th column. By means of (1) each coördinate can be expressed in the form

$$\begin{vmatrix} c_0\lambda_1 + c_1\lambda_2 & d_0\lambda_1 + d_1\lambda_2 \\ c_0\lambda_1' + c_1\lambda'_2 & d_0\lambda'_1 + d_1\lambda'_2 \end{vmatrix} = \begin{vmatrix} c_0 & c_1 \\ d_0 & d_1 \end{vmatrix} \cdot \begin{vmatrix} \lambda_1 & \lambda'_1 \\ \lambda_2 & \lambda'_2 \end{vmatrix},$$

hence $\lambda_1\lambda'_2 - \lambda_2\lambda'_1$ is removed as a factor from each v_i, and hence from the transformation.

The discriminant of (1) is quadratic in $\phi_1 : \phi_2$, hence there are two sets of k quadrics which appear once for all as factors. The coincident directrices are not fundamental in the transformation. The curve C_4 meets R in $4(m_1 + m_2)$ points, each of which is an actual k fold point on γ. These count as $2^k(k-1)(m_1 + m_2)$ actual double points, which must be added to the number of apparent double points of γ to determine the number of common bisecants of γ and C_4. The number of common bisecants u is $4p + 8$.

Let a common bisecant of C_4, γ meet the latter in P, P'. Through each

of these points passes a generator g, g' of R. Then g, g' are associated directrix lines associated with the quadric of the pencil on which PP' lies. Every common bisecant is a fundamental line of the second kind.

The number of parasitic conics is $p - k$, each appearing as a double base curve on the web of surfaces conjugate to the planes of space.

Each quadric of the pencil (1) is invariant under I. It contains C_4 simply but does not contain γ nor any parasitic curve.

4. Table of characteristics.

$$
\begin{aligned}
&S_1 \sim S_{4p+5} : {C_4}^{2p+1}\ {\gamma^2}_{2p+2+2k}(p-k){C_2}^2(4p+8)u\\
&C_4 \sim K_{8p+8} : {C_4}^{4p+1}\ {\gamma^4}_{2p+2+2k}(p-k){C_2}^4(4p+8)u^2\\
&\gamma \sim \Gamma_{8p+8} : {C_4}^{4p+2}\ {\gamma^3}_{2p+2+2k}(p-k){C_2}^4(4p+8)u^2\\
&\sigma \sim \sigma_{2p+4} : {C_4}^{p+1}\ \gamma_{2p+2+2k}(p-k)C_2(4p+8)u.\\
&J = K\Gamma.
\end{aligned}
$$

5. The complex *PP'*. The line joining a point P to its conjugate P' in I contains k such pairs P, P'. The equations are

$$(ax)(by) - (ay)(bx) = 0, \qquad (a'x)(b'y) - (a'y)(b'x) = 0,$$

which may be written in the form

$$\Sigma r_{ik}p_{ik} = 0, \qquad \Sigma r'_{ik}p_{ik} = 0, \qquad r = a_ib_k - a_kb_i, \qquad p_{ik} = x_iy_k - x_ky_i.$$

Each r_{ik} is a polynomial of degree $m_1 + m_2$ in λ_1, λ_2, r'_{ik} is the same polynomial in λ'_1, λ'_2. The equations may be written

$$c_0\lambda^{m_1+m_2} + \cdots + c_{m_1+m_2} = 0, \qquad c_0\lambda'^{m_1+m_2} + \cdots + c_{m_1+m_2} = 0,$$

in which each c_i is linear in p_{ik}, the Plücker coördinates of the line PP'. They may be replaced by two others, each symmetric in λ, λ'. Then by (1) these can be replaced by equations in ϕ_1/ϕ_2. The ϕ eliminant equated to zero is the equation of the complex. It is not a function of k. The order of the complex is $m_1 + m_2 - 1$. It can be linear only when $m_1 = m_2 = 1$.

Every point of R is singular on the complex. Let P be any point of R, g_1 the generator of R passing through it, and g_2 the conjugate of g_1 on R. The pencil P, g_2 belongs to the complex. Similarly, every tangent plane to R is a singular plane. Among the lines of the congruence g_1, g_2 there are two which also belong to another, say g'_1, g'_2. The common transversals of these

four lines are all double lines of the complex, each containing $2k$ pairs of conjugate points.

The congruence of double lines is self dual. Its order is equal to the number of double generators of a general complex cone. Since R and g'_2 are rational, this number is the maximum for a cone of its order.

If g_1, g_2 are fixed and g'_1, g'_2 describes R, the double lines generate a ruled surface having g_1, g_2 for $(m_1 + m_2)$-fold directrices; the whole congruence of double lines is arranged on a system of these ruled surfaces. If R has a rectilinear directrix, it belongs to every ruled surface of the system; if R has two rectilinear directrices, the entire congruence of double lines is replaced by these two directrix lines. Finally, if $m_1 = m_2 = 1$, the congruence is replaced by the other regulus of R. Since each generator of this regulus contains an infinite number of pairs of conjugate points, it remains invariant under I, hence R is transformed into itself.

6. Earlier particular cases. All cases already found in the literature are included in the present paper. The case in which the two directrices are fixed belongs to the types discussed by de Paolis.[1] It can be reduced to the monoidal type. Montesano[2] considered the case $m_1 = m_2 = k = 1$ from a different stand-point. This was generalized for any k by Snyder.[3] These two are contained in a general linear complex. Various cases in which the complex PP' is composed of secants to a rational curve are given by Davis,[4] Black,[5] Carroll,[6] Dye,[7] Montague,[8] Vicianze,[9] and Schoonmaker and Snyder.[10] A decidedly more general type is given by Dye.[11]

---

[1] *Rom. Acc. Lin. Mem.* (4), vol. 1 (1885), pp. 576-608; *Rom. Acc. Lin. Rend.* (4), vol. 1 (1885), pp. 735-792, 754-758.

[2] *Rom. Acc. L. Rend.* (4), vol. 4 (1888), pp. 207-25, 277-285.

[3] *Atti Congresso intern. di Bologna*, vol. 4 (1928), pp. 13-20.

[4] *American Journal of Mathematics*, vol. 52 (1930), pp. 53-71; vol. 53 (1931), pp. 72-80; *Tôhoku Mathematical Journal*, vol. 33 (1931), pp. 254-259.

[5] *Transactions of the American Mathematical Society*, vol. 34 (1931), pp. 795-810; *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 417-420; vol. 42 (1936), pp. 535-540.

[6] *American Journal of Mathematics*, vol. 54 (1931), pp. 707-717; vol. 56 (1934), pp. 96-108.

[7] *Transactions of the American Mathematical Society*, vol. 32 (1930), pp. 251-261; *American Journal of Mathematics*, vol. 54 (1932), pp. 499-504 (with F. R. Sharpe).

[8] *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 727-731.

[9] *Dissertation*, Catania University (1922).

[10] *American Journal of Mathematics*, vol. 54 (1932), pp. 299-304.

[11] *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 535-540, and a paper presented at the Duke meeting, December, 1936.

7. Extension to S_n. If in (3) each polynomial ψ_i is of degree $n-1$ in λ_1, λ_2, while in (1) and (2) the number of homogeneous variables x_i is $n+1$, the transformation I is at once extended to S_n. Instead of two skew directrix lines there are now $n-1$ linear spaces S_{n-2} defined by (2) and (3), forming an involution of order $n-1$ among the generating S_{n-2}'s of R.

Through an arbitrary point y can be drawn just one transversal t which meets all $n-1$ associated director spaces.

The fundamental elements C_4 and γ are defined as in S_3; C_4 is the base of the pencil of quadric primals, and γ is the locus of the intersection of a director S_{n-2} with its associated set of k quadric primals. The image of each point of γ is a conic cut from the primals by the plane meeting the other director spaces.

In S_n, the line through a point P on γ, in order to be parasitic must not only be a bisecant of C_4, hence a line of the quadric primal through P, but must also meet all the other director spaces S_{n-2} in points of their quadrics of intersection with their associated primals. Since every point of γ is double on $|S|$, these lines have $n-1$ double points and two points of multiplicity $2p+1$ on C_4, and are therefore base lines of the system.

The image of C_4 is a primal K, and of γ is a primal Γ. These two primals constitute the complete Jacobian of the system.

As in S_3, those values of $\mu_1, \mu_2 = \mu$ corresponding to double roots of (3) define primals of the pencil (1) which are factors of the system, and therefore do not appear as fundamental elements in the proper transformation.

In S_3 the phenomenon of a plane field of lines meeting both directrices is possible only when the directrices intersect.

In S_4, associated with every value of μ each triad of director planes determines a plane, that through the three points in which they meet by twos. This plane meets each director plane in a line, hence every line in it meets all three director planes. The plane meets its associated quadric primals in k conics, all parasitic in I. Thus, in S_4, I has ∞^1 fundamental conics of the second kind, forming a surface Σ, which is a part of the base of $|S|$, conjugate to the primes of S_4.

In S_n there are $n-1$ spaces S_{n-2} meeting by twos in $\binom{n-1}{2}$ spaces S_{n-4}. Through any point of any one of these intersections can be drawn one line meeting the others in each of the director spaces through it. The plane of these two intersecting lines meets each director space in a line. Hence for each set of k quadrics of the pencil there are ∞^{n-4} such parasitic conics,

generating a manifold Σ of $n-2$ dimensions, not lying in a linear space of dimensions less than n. It is that defined by the Veneroni transformation.[12]

These parasitic conics include all the particular cases arising from base elements lying in simpler spaces for special values of μ.

Finally, the other base elements of $|S|$ consist of the lines that are bisecants of the quartic C_4 and multisecants of γ. They form a manifold T.

The equations of the transversal line t through y meeting the director spaces are $\sum_{i=1}^{n+1} P_i^{(j)} x_i = 0 \quad (j = 1, 2, \cdots, n-1)$. The point z in which t meets the prime $\Sigma x_i = 0$ is given by the matrix

$$\begin{Vmatrix} P_1^{(1)} & \cdots & P_{n+1}^{(1)} \\ \cdots & \cdots & \cdots \\ P_1^{(n-1)} & \cdots & P_{n+1}^{(n-1)} \\ 1 & \cdots & 1 \end{Vmatrix}.$$

By (3), each element can be reduced to a polynomial r_{ik} of degree $n-2$ in λ, the coefficients being linear in y and functions of order $(m_1 + m_2) - (n-2)$ in ϕ_1, ϕ_2. Each coördinate z_i is expressed as a determinant of order $n-1$ with elements r_{ik}. But this is the product of two determinants, one in the coefficients, and the other is $|\lambda_1^{n-2}, \lambda_2^{n-1}, \cdots, 1| = \pi(\lambda_i - \lambda_k)$.

This second factor is the same for all z_i, and divides out of the transformation. This explains why the coincident director spaces do not appear as fundamental elements in I.

A more convenient point z is obtained by adding the points of intersection of t with the director spaces. Each z_i is of order $(m_1 + m_2 - (n-2))$ in ϕ_1, ϕ_2 and contains y to order $n-2$ apart from ϕ. Hence the order of the transformation I is

$$4k(m_1 + m_2 - n + 2) + 2(n-2) + 3.$$

Let $p = k(m_1 + m_2 - (n-2))$. Then the order of I is $4p + 2n - 1$. The base C_4 appears to multiplicity $2p + 1$. The manifold γ of $n-2$ dimensions is double; the image of each point on it is a conic. By eliminating λ between (2) and (3) and making use of (1), the order of γ is found to be $r = 2p + 2k(n-2) + n - 1$. The order of Σ, locus of the parasitic conics may be obtained by determining the plane associated with each point

[12] Snyder and Rusk, "The Veneroni transformation in S_n," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 585-592.

of intersection of two director spaces, and getting the k conics cut from the associated quadric primals by this plane. This set of conics, as the point describes the intersection of the director spaces, describes Σ. It is of order $(n-2)(4p+n-1)-r$. The manifold of bisecants of C_4 and multisecants of γ is of order $4p+6n-10$. Since each quadric primal of the pencil goes into itself by I, the order of K is at once found to be $8p+4n-4$, hence that of Γ is $(n-1)(4p+2n-2)$. The table of characteristics for primes and the fundamental elements is now complete.

$$
\begin{aligned}
S_1 &\sim S_{4p+2n-1} : {C_4}^{2p+1}\, {\gamma_r}^2\, \Sigma^2 T \\
C_4 &\sim K_{8p+4n-2} : {C_4}^{4p+1}\, {\gamma_r}^4\, \Sigma^2 T^2 \\
\gamma &\sim \Gamma_{(n-1)(4p+2n-2)} : {C_4}^{4p+2}\, {\gamma_r}^3\, \Sigma^4 T^2 \\
\sigma &\sim \sigma_{2p+n+1} : {C_4}^{p+1}\quad \gamma_r \qquad\quad \Sigma\, T \\
J &= K\Gamma.
\end{aligned}
$$

CORNELL UNIVERSITY,
AND
WELLS COLLEGE.

THE VENERONI TRANSFORMATION IN S_n.*

By GERTRUDE BLANCH.

1. Introduction. The primals of S_n, of order n, which pass through $(n+1)$ general linear spaces S_{n-2} form a homaloidal system. The Cremona transformation of order n, determined by this system is known as the *Veneroni*[1] transformation. Studies of it have been made by Eiesland[2] and more recently, by J. A. Todd[3] and by Virgil Snyder.[4] The author[5] studied the transformation in S_4; but the form used does not permit of a ready generalization for S_n. One general form has been obtained by Snyder and Rusk.[6] The form derived in this paper may be of interest because of its special symmetry. In addition, it is shown that in S_n, $n > 3$, it is necessary to satisfy $n(n-3)/2 - 1$ conditions in order that the transformation shall be involutorial.

2. Following C. A. Rupp,[7] let the linear forms S_{n-2} be defined by:

$$\phi_i: \quad x_i = 0; \quad A_i(x) = \begin{cases} \sum_{j=1}^{n+1} a_{ij}x_j = 0; & (i = 1, 2, \cdots, (n+1) \\ a_{ii} = 0. \end{cases}$$

Consider the determinant of order $(n+1)$ given below:

$$(1) \qquad \begin{vmatrix} k_1 & a_{12}x_2 & a_{13}x_3 & \cdots & a_{1r}x_r & \cdots & a_{1,n+1}x_{n+1} \\ k_2 & -A_2(x) & a_{23}x_3 & \cdots & a_{2r}x_r & \cdots & a_{2,n+1}x_{n+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ k_r & a_{r2}x_2 & a_{r3}x_3 & \cdots & -A_r(x) & \cdots & a_{r,n+1}x_{n+1} \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdots & \cdot \\ k_{n+1} & a_{n+1,2}x_2 & a_{n+1,3}x_3 & \cdots & a_{n+1,r}x_r & \cdots & -A_{n+1}(x) \end{vmatrix} = 0.$$

* Received August 19, 1936; revised May 3, 1937.

[1] *Lombardo Rendiconti* II, vol. 34 (1901), pp. 640-654.

[2] *Palermo Rendiconti*, vol. 54 (1930), pp. 335-365.

[3] *Proceedings of the Cambridge Philosophical Society*, vol. 26 (1930), pp. 323-333.

[4] *Bulletin of the American Mathematical Society*, vol. 40 (1934), pp. 673-687.

[5] *American Journal of Mathematics*, vol. 58 (1936), pp. 639-645.

[6] *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 585-592.

[7] *Bulletin of the American Mathematical Society*, vol. 35 (1929), pp. 319-320.

The cofactor of every k_j is satisfied by $x_i = A_i(x) = 0$ $(i = 2, 3 \cdots n+1)$. Moreover, the cofactor of every k_j is also satisfied by $x_1 = A_1(x) = 0$. For add to the second column the elements of all subsequent columns. The elements of the second column then become, respectively,

$$A_1(x), -a_{21}x_1, \cdots, -a_{r1}x_1, \text{ etc.}$$

It can be readily verified that the cofactor of k_i is $x_i W_i$, where W_i is a V^{n-1}_{n-1}, containing all the fundamental S_{n-2} except ϕ_i. Hence the determinant (1) is the complete linear system $| V_n^n |$, with k_i arbitrary. Let

$$x'_i = x_i W_i \qquad (i = 1, 2, \cdots, n+1). \tag{2}$$

Replacing the first column of (1) by the i-th column, we get a determinant with two columns alike, and on account of (2), we have the following identities:

$$M_i: \quad x_i \sum_j a_{ji} x'_j - A_i(x) x'_i = 0 \qquad (i = 2, \cdots, n+1).$$

Adding $M_2 + M_3 + \cdots M_{n+1}$, and remembering that $A_i(x) = \sum_j a_{ij} x_j$, $a_{ii} = 0$, we get

$$M_1: \quad x_1 \sum_j a_{j1} x'_1 - A_1(x) x'_1 = 0.$$

We therefore have the $(n+1)$ bilinear equations:

$$x_i \sum_j a_{ji} x'_j - A_i(x) x'_i = 0 \qquad (i = 1 \cdots n+1). \tag{3}$$

Write the matrix of the coefficients $\| a_{ij} \|$. It should be observed that the elements of the i-th row of this matrix determine the fundamental ϕ_i in (x) and similarly, the transpose of this same matrix gives at once the form of the fundamental S_{n-2} of (x'), namely:

$$\phi'_i: \quad x'_i = \sum_j a_{ji} x'_j = 0 \qquad (i = 1 \cdots n+1). \tag{4}$$

3. Involutorial transformations. If $a_{ij} = a_{ji}$ for every i, j, and (x), (x') are regarded as superposed, we have the identity transformation; if $a_{ij} = -a_{ji}$ for every i, j, then the transformation becomes a polarity among $(n+1)$ quadric primals, and is necessarily involutorial. However, by subjecting (x') to a linear transformation—equivalent to changing the frame of reference—we may obtain a form wherein the fundamental S_{n-2} in (x') coincide with their *associated*[8] S_{n-2} in (x), with fewer restrictions.

[8] By *associated* forms we mean those of (x) and (x') associated in the same bilinear equation.

To determine a suitable transformation carrying $\phi'_i \sim \phi_i$, we note from (4) that x'_i will have to be some linear combination of the two primes determining ϕ_i. Hence the transformation must be of the form

$$x'_i = r_i y_i + s_i A_i(y), \qquad (i = 1 \cdots n+1). \tag{5}$$

As a consequence of (5)

$$\sum_j a_{ji} x'_j \sim \sum_j a_{ji}(r_j y_j + s_j \sum_{k=1}^{n+1} a_{jk} y_k) = \mu_i(y). \tag{6}$$

It is necessary that $\mu_i(y)$ also be of the form

$$\rho_i y_i + t_i A_i(y). \tag{7}$$

From (6) and (7)

$$(\rho_i - \sum_j s_j a_{ji}^2) y_i + \sum_k (t_i a_{ik} - a_{ki} r_k - \sum_j s_j a_{ji} a_{jk}) y_k \equiv 0, \quad k \neq i. \tag{8}$$

The coefficient of every y_j must vanish. Consequently set $\rho_i = \sum_j s_j a_{ji}^2$. As to the second group of terms, we note that when $j = k$, $t_i a_{ik} - r_k a_{ki} = 0$, since $a_{kk} = 0$. Set $i = 1$, $k = 2$; and again $i = 2$, $k = 1$. We thus get

$$t_1 a_{12} - a_{21} r_2 = 0; \qquad t_2 a_{21} - a_{12} r_1 = 0.$$

Or, transposing,

$$a_{12}(t_1 + r_1) = a_{21}(t_2 + r_2).$$

Similarly, we must satisfy, in general

$$a_{ij}(t_i + r_i) = a_{ji}(t_j + r_j); \text{ for } (i, j = 1 \cdots (n+1).$$

Since the object is to impose as few restrictions as possible on the coefficients a_{ij}, set

$$t_i = -r_i; \qquad (i = 1 \cdots (n+1). \tag{9}$$

From (8) and (9)

$$r_i a_{ik} + r_k a_{ki} + \sum_j s_j a_{ji} a_{jk} = 0; \qquad (i, k = 1 \cdots (n+1); \; k \neq i). \tag{10}$$

By assigning all possible values to i and k, we obtain $n(n+1)/2$ linear equations in the $2(n+1)$ variables r_k, s_k. A study of the forms in S_5 and generalization to S_n shows that these equations determine $n(n-3)/2 - 1$ functionally independent determinants in the space of the coefficients a_{ij}, which must have the value zero if the conditions (10) are to be satisfied.[9]

[9] Thus in S_5 the 15×12 matrix gives 4 determinants of order 12, all having the last eleven rows of the matrix in common. The matrix of the last eleven rows has rank

When (10) holds, the forms are necessarily involutorial. For then the bilinear equations (3) are transformed into:

$$(11) \qquad x_i(\rho_i y_i - r_i \sum_j a_{ij} y_j) - \sum_j a_{ij} x_j \cdot (r_i y_i + s_i \sum_k a_{ik} y_k) = 0.$$

It is readily verified that the coefficient of $x_p y_q$ is the same as that of $x_q y_p$ for all subscripts. Hence the theorem: *The Veneroni Transformation in S_n can be made involutorial, with coincident associated fundamental elements, by imposing $n(n-3)/2 - 1$ conditions on the coefficients.* This verifies that in S_3 the transformation may always be made involutorial, and indeed in ∞' ways; in S_4, one condition must be imposed on the coefficients—as the author verified in the paper referred to. In S_5, four conditions among the coefficients of the form are necessary.

These involutorial transformations are not the only ones obtainable—the S_{n-2} of (x') need not all be associated with the same spaces of (x) in the bilinear equations. But the author's study in S_4 previously mentioned shows that other such involutorial transformations are always more specialized than the type where the S_{n-2} of (x') are the same as their associated S_{n-2} of (x).

HUNTER COLLEGE,
NEW YORK.

11; hence by well-known theorems about such matrices, the vanishing of these 4 determinants insures the vanishing of every other determinant of order 12. It is further possible to find sets of values of the coefficients a_{ij} for which the Jacobian of the four determinants with respect to certain four of the coefficients a_{ij} does not vanish. Moreover, for the most general values of these coefficients, the four determinants have no common factor. Consequently the 4 conditions are independent in S_5. The form of the determinants in S_5 permits a ready extension by induction to the independence of the conditions in S_n.

ON $2n$ POINTS WITH A REAL CROSS-RATIO.*

By F. MORLEY and J. R. MUSSELMAN.*

1. The viewpoint of Klein, that elementary geometry is not a study of isolated theorems but the theory of invariants of a group, is applied here to the inversive group and we study the meaning of an invariant of that group. Let us choose any $2n$ points $a_i (i = 1, 2, \cdots, 2n)$ in a plane (or on a sphere) which are ordered. Then, the expression

$$\frac{(a_1 - a_2)(a_3 - a_4) \cdots (a_{2n-1} - a_{2n})}{(a_2 - a_3)(a_4 - a_5) \cdots (a_{2n} - a_1)} = \rho, \tag{1.1}$$

where ρ is any real number, is invariant under all homographies

$$y = (\alpha x + \beta)/(\gamma x + \delta)$$

and under all antigraphies

$$\bar{y} = (\alpha x + \beta)/(\gamma x + \delta),$$

and hence is an invariant of the inversive group.

2. For $n = 2$, the cross-ratio

$$\frac{(a_1 - a_2)(a_3 - a_4)}{(a_2 - a_3)(a_4 - a_1)} = \rho \tag{2.1}$$

is the necessary and sufficient condition that *the four points a_i lie on a circle.*

For $n = 3$, we have six points a_i and

$$\frac{(a_1 - a_2)(a_3 - a_4)(a_5 - a_6)}{(a_2 - a_3)(a_4 - a_5)(a_6 - a_1)} = \rho. \tag{2.2}$$

Consider the three arcs $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_1$ (or else the arcs $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_1a_2$). Let $a_1a_2a_3$ and $a_3a_4a_5$ meet again at a point b_3. Then from (2.1) we have

$$\frac{(a_1 - a_2)(a_3 - b_3)}{(a_2 - a_3)(b_3 - a_1)} = \rho_1 \quad \text{and} \quad \frac{(a_3 - a_4)(a_5 - b_3)}{(a_4 - a_5)(b_3 - a_3)} = \rho_2.$$

* Received April 30, 1937.

Hence, by using (2. 2) we have

$$\frac{(a_5 - a_6)(a_1 - b_3)}{(a_6 - a_1)(b_3 - a_5)} = \rho_3.$$

That is, the three arcs $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_1$ meet at the point b_3. And equally the three arcs $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_1a_2$ meet at a point, say b_4. The single condition (2. 2) implies both facts. Two arcs $a_1a_2a_3$, $a_3a_4a_5$ which meet again at b_3 have a resultant arc $a_1b_3a_5$. The arcs $a_1a_2a_3$, $a_3a_4a_5$ and $a_5b_3a_1$ are then termed *balanced*. Thus for an ordered hexagon the invariant (2. 2) means that *the arcs* $a_1a_2a_3$, $a_3a_4a_5$ and $a_5a_6a_1$ (or $a_2a_3a_4$, $a_4a_5a_6$ and $a_6a_1a_2$) *are balanced.*

For $n = 4$, we have eight points and

$$(2.3) \qquad \frac{(a_1 - a_2)(a_3 - a_4)(a_5 - a_6)(a_7 - a_8)}{(a_2 - a_3)(a_4 - a_5)(a_6 - a_7)(a_8 - a_1)} = \rho.$$

Consider the four arcs $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_7$ and $a_7a_8a_1$ (or else the arcs $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_7a_8$ and $a_8a_1a_2$). Let $a_1a_2a_3$ and $a_3a_4a_5$ meet again at b_3, and let $a_5a_6a_7$ and $a_7a_8a_1$ meet again at b_7. Then from (2. 1) we have

$$\frac{(a_1 - a_2)(a_3 - b_3)}{(a_2 - a_3)(b_3 - a_1)} = \rho_1, \qquad \frac{(a_3 - a_4)(a_5 - b_3)}{(a_4 - a_5)(b_3 - a_3)} = \rho_2,$$

$$\frac{(a_5 - a_6)(a_7 - b_7)}{(a_6 - a_7)(b_7 - a_5)} = \rho_3, \qquad \frac{(a_7 - a_8)(a_1 - b_7)}{(a_8 - a_1)(b_7 - a_7)} = \rho_4.$$

Hence, by using (2. 3), we have

$$\frac{(a_5 - b_3)(a_1 - b_7)}{(b_3 - a_1)(b_7 - a_5)} = \rho_5,$$

which implies that the points a_5, b_3, a_1 and b_7 are concyclic. Thus for an ordered octagon the invariant (2. 3) is satisfied *if the four points* a_1, a_5, *the second meet of* $a_1a_2a_3$ *and* $a_3a_4a_5$, *and the second meet of* $a_5a_6a_7$ *and* $a_7a_8a_1$ *are concyclic.*

This can also be stated as follows: *if we choose any two points* x *and* y *concyclic with* a_1 *and* a_5 *such that the two sets of arcs* $a_1a_2a_3$, $a_3a_4a_5$, a_5xa_1 *and* $a_5a_6a_7$, $a_7a_8a_1$, a_1ya_5 *are balanced, the invariant* (2. 3) *is satisfied.* For this implies that $a_5b_3a_1b_7$ are concyclic.

If we call the second intersection of the arcs $a_{i-2}a_{i-1}a_i$ and $a_ia_{i+1}a_{i+2}$ by b_i (all subscripts taken modulo 8) we have the theorem that *if the four points* a_5, b_3, a_1 and b_7 *are concyclic, so also are* a_6, b_4, a_2, b_8; a_7, b_5, a_3, b_1 *and* a_8, b_6, a_4, b_2.

For the condition from (2. 1) that a_{i+1}, b_{i-1}, a_{i-3}, b_{i+3} be concyclic reduces to

$$\frac{(a_{i-3}-a_{i-2})(a_{i-1}-a_i)(a_{i+1}-a_{i+2})(a_{i+3}-a_{i+4})}{(a_{i-2}-a_{i-1})(a_i-a_{i+1})(a_{i+2}-a_{i+3})(a_{i+4}-a_{i-3})}=\rho. \tag{2.4}$$

If we let $i=4$ or 6 in (2. 4) we obtain (2. 3); if we let $i=5$ or 7 in (2. 4) we obtain the reciprocal of (2. 3) which proves the theorem. Thus if it happens that one set of points a_{i+1}, b_{i-1}, a_{i-3}, b_{i+3} for an ordered octagon be concyclic, there are three other concyclic sets. The single condition (2. 3) implies all four.

In the general case of $2n$ ordered points $a_1a_2a_3a_4\cdots a_{2n}$, we replace the arcs $a_1a_2a_3$, $a_3a_4a_5$ by the arc $a_1b_3a_5$ where b_3 is the other intersection of the two arcs. We have now $2n-2$ points

$$a_1b_3a_5a_6a_7\cdots.$$

We replace the arcs $a_1b_3a_5$, $a_5a_6a_7$ by the arc $a_1c_5a_7$ where c_5 is the second intersection of the two arcs $a_1b_3a_5$ and $a_5a_6a_7$. We have now $2n-4$ points

$$a_1c_5a_7a_8a_9a_{10}\cdots.$$

Proceeding in this manner we arrive at four points. If these are concyclic, then the original set (or any intermediate set) has a real cross-ratio. For we have taken b_3 such that

$$\frac{(a_1-a_2)(a_3-b_3)}{(a_2-a_3)(b_3-a_1)}=\rho_1 \quad \text{and} \quad \frac{(a_3-a_4)(a_5-b_3)}{(a_4-a_5)(b_3-a_3)}=\rho_2$$

or

$$\frac{(a_1-a_2)(a_3-a_4)}{(a_2-a_3)(a_4-a_5)}=\rho_1\rho_2 \quad \frac{a_1-b_3}{a_5-b_3}$$

so that instead of

$$\frac{(a_1-a_2)(a_3-a_4)(a_5-a_6)\cdots}{(a_2-a_3)(a_4-a_5)(a_6-a_7)\cdots}=\rho$$

we have

$$\frac{(a_1-b_3)(a_5-a_6)\cdots}{(b_3-a_5)(a_6-a_7)\cdots}=\rho'.$$

To sum up: if we have a closed arc polygon $a_1a_2a_3$, $a_3a_4a_5$, $\cdots$, $a_{2n-1}a_{2n}a_1$ it is balanced if upon replacing $a_1a_2a_3$, $a_3a_4a_5$ by the resultant arc $a_1b_3a_5$, the new polygon is balanced. That is, if by continuing the process we get four points on a circle. If we have such a polygon, then taking $a_2, a_4\cdots a_{2n}$

arbitrarily on their arcs we get $2n$ points with a real cross-ratio. In particular if a_2 approaches a_1, if a_4 approaches a_3, and so on, then

$$\frac{\delta a_1 \cdot \delta a_3 \cdot \cdots \delta_{2n-1}}{(a_1 - a_3)(a_3 - a_5) \cdots (a_{2n-1} - a_1)} \text{ is a real} = \delta\rho.$$

For $n = 6$ we can give another interpretation of

$$(2.5) \qquad \frac{(a_1 - a_2)(a_3 - a_4) \cdots (a_{11} - a_{12})}{(a_2 - a_3)(a_4 - a_5) \cdots (a_{12} - a_1)} = \rho.$$

Let the second intersection of arcs $a_{i-2}a_{i-1}a_i$ and $a_ia_{i+1}a_{i+2}$ be denoted by b_i. We shall derive the condition that the three circles $a_{i-6}b_{i-4}a_{i-2}$, $a_{i-2}b_ia_{i+2}$ and $a_{i+2}b_{i+4}a_{i-6}$ meet at a point. The condition that b_{i-4} is the second intersection of $a_{i-6}a_{i-5}a_{i-4}$ and $a_{i-4}a_{i-3}a_{i-2}$ is

$$\frac{(a_{i-6} - a_{i-5})(a_{i-4} - a_{i-3})(a_{i-2} - b_{i-4})}{(a_{i-5} - a_{i-4})(a_{i-3} - a_{i-2})(b_{i-4} - a_{i-6})} = \rho_1.$$

The condition that b_i is the second intersection of $a_{i-2}a_{i-2}a_i$ and $a_ia_{i+1}a_{i+2}$ is

$$\frac{(a_{i-2} - a_{i-1})(a_i - a_{i+1})(a_{i+2} - b_i)}{(a_{i-1} - a_i)(a_{i+1} - a_{i+2})(b_i - a_{i-2})} = \rho_2.$$

The condition that b_{i+4} is the second intersection of $a_{i+2}a_{i+3}a_{i+4}$ and $a_{i+4}a_{i+5}a_{i-6}$ is

$$\frac{(a_{i+2} - a_{i+3})(a_{i+4} - a_{i+5})(a_{i-6} - b_{i+4})}{(a_{i+3} - a_{i+4})(a_{i+5} - a_{i-6})(b_{i+4} - a_{i+2})} = \rho_3.$$

The condition that circles $a_{i-6}b_{i-4}a_{i-2}$, $a_{i-2}b_ia_{i+2}$, $a_{i+2}b_{i+4}a_{i-6}$ meet at a point is from (2.2)

$$\frac{(a_{i-6} - b_{i-4})(a_{i-2} - b_i)(a_{i+2} - b_{i+4})}{(b_{i-4} - a_{i-2})(b_i - a_{i+2})(b_{i+4} - a_{i-6})} = \rho_4.$$

The product of these four expressions is

$$\frac{(a_{i-6} - a_{i-5})(a_{i-4} - a_{i-3})(a_{i-2} - a_{i-1})(a_i - a_{i+1})(a_{i+2} - a_{i+3})(a_{i+4} - a_{i+5})}{(a_{i-5} - a_{i-4})(a_{i-3} - a_{i-2})(a_{i-1} - a_i)(a_{i+1} - a_{i+2})(a_{i+3} - a_{i+4})(a_{i+5} - a_{i-6})} = \rho.$$

For $i = 7$ or 9—subscripts are taken modulo 12—we get (2.5); for $i = 8$ or 10 we get the reciprocal of (2.5). Other values of i, due to the symmetry, repeat. Hence the invariant (2.5) for an ordered twelve-point means that *the circles* $a_1b_3a_5$, $a_5b_7a_9$, $a_9b_{11}a_1$ *are on a point.* Equally it means that circles $a_2b_4a_6$, $a_6b_8a_{10}$, $a_{10}b_{12}a_2$; $a_3b_5a_7$, $a_7b_9a_{11}$, $a_{11}b_1a_3$; $a_4b_6a_8$, $a_8b_{10}a_{12}$, $a_{12}b_2a_4$ are on a point. The single condition (2.5) implies all four facts.

In similar fashion one can prove for sixteen points that if we call the

second intersection of $a_{i-2}a_{i-1}a_i$ and $a_ia_{i+1}a_{i+2}$ by b_i and the second intersection of $a_{i-4}b_{i-2}a_i$ and $a_ib_{i+2}a_{i+4}$ by c_i then the reality of the cross-ratio (1.1) implies that $a_{i-6}c_{i-2}a_{i+2}c_{i+6}$ are concyclic. There are eight such sets of four points, such if one set is concyclic, all eight are. We list them

$$\begin{array}{ll} a_1c_5a_9\ c_{13} & a_5c_9\ a_{13}c_1 \\ a_2c_6a_{10}c_{14} & a_6c_{10}a_{14}c_2 \\ a_3c_7a_{11}c_{15} & a_7c_{11}a_{15}c_3 \\ a_4c_8a_{12}c_{16} & a_8c_{12}a_{16}c_4 \end{array}$$

For ten points the invariant (1.1) implies that *the circles* $a_{i-4}b_{i-2}a_i$ *and* $a_ib_{i+2}a_{i+4}$ *have* a_{i+5} *as their second point of intersection.* If it happens once it happens for all ten pairs of circles.

For fourteen points, we can state that if the invariant (1.1) be satisfied then *circles* $a_{i-8}c_{i-4}a_i$ *and* $a_ib_{i+2}a_{i+4}$ *have* a_{i+5} *as their second point of intersection.* If it happens once, it happens for all fourteen sets of circles.

3. A similar argument applies when the cross-ratio of the $2n$ points is a pure imaginary ρi. The circle on abc can here be replaced by the Apollonian circle on b and about ac. Let us investigate six ordered points where

$$\frac{(a_1 - a_2)(a_3 - a_4)(a_5 - a_6)}{(a_2 - a_3)(a_4 - a_5)(a_6 - a_1)} = \rho i. \tag{3.1}$$

The Apollonian circle on a_2 and about a_1a_3 may be written

$$\frac{(a_1 - a_2)(a_3 - x)}{(a_3 - a_2)(a_1 - x)} = \rho_1 i.$$

Similarly the Apollonian circles on a_4 and about a_3a_5, on a_6 and about a_5a_1 are

$$\frac{(a_3 - a_4)(a_5 - x)}{(a_5 - a_4)(a_3 - x)} = \rho_2 i$$

$$\frac{(a_5 - a_6)(a_1 - x)}{(a_1 - a_6)(a_5 - x)} = \rho_3 i.$$

The product of these three expressions is (3.1). Hence for *an ordered hexagon the invariant* (3.1) *means that the Apollonian circles on* a_2 *and about* a_1a_3, *on* a_4 *and about* a_3a_5, *on* a_6 *and about* a_5a_1 *meet at a point.* Equally the Apollonian circles on a_3 and about a_2a_4, on a_5 and about a_4a_6, on a_1 and about a_6a_2 will meet at a point. The single condition (3.1) implies both facts.

4. We can express the condition that any n ordered points should lie

on a line in terms of the Lagrange resolvents [1] of the n-point. We term as Lagrange resolvents of an ordered n-point the $n-1$ expressions

$$V_i = a_1 + \epsilon a_2 + \epsilon^2 a_3 \cdots \epsilon^{n-1} a_n$$

where ϵ is a root of $\epsilon^n = 1$ other than $\epsilon = 1$. We can take the n points as $a_i = a + \rho_i b$, $(i = 1, \cdots, n)$ where the ρ_i are real. Consider for simplicity four points. Then

$$V_1 = b[(\rho_1 - \rho_3) + i(\rho_2 - \rho_4)]$$
$$V_2 = b[\rho_1 - \rho_2 + \rho_3 - \rho_4]$$
$$V_3 = b[(\rho_1 - \rho_3) - i(\rho_2 - \rho_4)]$$

whence

$$V_1\bar{V}_1 = V_3\bar{V}_3, \quad V_2\bar{V}_1 = V_3\bar{V}_2 \quad \text{and} \quad V_2\bar{V}_3 = V_1\bar{V}_2$$

or

$$V_1/\bar{V}_3 = V_2/\bar{V}_2 = V_3/\bar{V}_1.$$

And so in general, for n ordered points to be on a line, we have

$$V_1/\bar{V}_{n-1} = V_2/\bar{V}_{n-2} = \cdots = V_{n-1}/\bar{V}_1.$$

THE JOHNS HOPKINS UNIVERSITY,
AND
WESTERN RESERVE UNIVERSITY.

[1] The Morleys, *Inversive Geometry*, p. 203.

EXTRACT FROM A LETTER BY E. CARTAN CONCERNING MY NOTE: ON CLOSED SPACES OF CONSTANT MEAN CURVATURE.*

By T. Y. Thomas.

In the proof of the theorem in my recent note in the *American Journal of Mathematics* (Vol. LVIII, 1936, p. 702) on hypersurfaces of a Euclidean space the assumption was made that the hypersurfaces were closed (but without boundary). I am in receipt of a letter from Professor E. Cartan following a mathematical discussion at Princeton in which he shows that this theorem can be proved without the above assumption, i. e. the theorem is in reality of local character. Cartan's proof forms an interesting supplement to my previous note. I present it (with permission) as an extract taken verbatim from his letter:

"En ce qui concerne votre intéressant théorème, dont vous m'aviez du reste parlé à Princeton, sur les espaces à courbure moyenne constante, je ne sais pas si je trompe, mais il me semble qu'il se ramène à une propriété purement locale et qu'il n'est pas necéssaire de supposer l'espace clos. Si j'ai bien compris, il s'agit de démontrer qu'une hypersurface S à $n \geqq 3$ dimensions de l'espace euclidean à $n+1$ dimensions est à courbure constants si, considérée comme espace riemannien elle satisfait aux conditions $B_{\alpha\beta} = \lambda g_{\alpha\beta}$ avec $\lambda > 0$.

"Soient en effet $a_1, a_2, \cdots, a_n$ les courbures principales de l'hypersurface S. Les conditions imposées sont

$$\begin{aligned}
a_1a_2 + a_1a_3 + \cdots + a_1a_n &= \lambda,\\
a_2a_1 + a_2a_3 + \cdots + a_2a_n &= \lambda,\\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot &\\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot &\\
a_na_1 + a_na_2 + \cdots + a_na_n &= \lambda,
\end{aligned}$$

on encore, un désignant par S les somme des courbures,

$$a_i^2 - a_iS + \lambda = 0.$$

"Toutes les courbures principales satisfont donc à une même equation du second degré

$$x^2 - xS + \lambda = 0.$$

"Supposons qu'elles ne soient pas toutes égales entre elles; et que par exemple $a_1 \neq a_2$. Toutes les autres courbures seront égales soit à a_1, soit à a_2. Supposons que a_1 se présente p fois, et a_2 q fois $(p + q = n)$. On aura alors

* Received July 12, 1937.

$$a_1 + a_2 = S, \qquad a_1 a_2 = \lambda,$$

et, par suite,

$$(p-1)a_1 + (q-1)a_2 = 0, \qquad a_1 a_2 > 0.$$

Ces deux conditions sont contradictoires.

"Toutes les courbures principales étant égales, l'hypersurface est une hypersphere de rayon $1/|\, a_1 \,|$."

It suffices for the above demonstration that the hypersurface S is defined by functions $\phi^i(x)$ where $i = 1, \cdots, n+1$ which are continuous and possess continuous partial derivatives to the third order (i. e. S is a hypersurface of class C^3). Including the case $n = 2$ for which the above result is immediately valid we may state the following theorem:

Any hypersurface of class C^3 of constant mean curvature $\lambda > 0$ and dimensionality $n \geqq 2$ in a Euclidean space of $n+1$ dimensions is a space of constant curvature.

Remark. One can find n mutually orthogonal directions $\lambda_1, \cdots, \lambda_n$ at a point P of the hypersurface, these directions being given as solutions of the equations $\Sigma_j (b_{ij} - a_k g_{ij})\lambda_k{}^j = 0$ where the b_{ij} are the coefficients of the second fundamental form and the a_k are the (real) roots of the determinant equation $|\, b_{ij} - a g_{ij} \,| = 0$. The $\lambda_1, \cdots, \lambda_n$ define the directions of lines of curvature at P and the $a_1, \cdots, a_n$ are (by definition) the normal curvatures at P of the corresponding lines of curvature. Denote by r_{ij} the Gaussian curvature at P of the geodesic surface determined by the directions λ_i and λ_j $(i \neq j)$. Then $r_{ij} = a_i a_j$ (Eisenhart, *Riemannian Geometry*, 1926, p. 198). Now $\Sigma_j r_{ij} = \lambda$ $(i \neq j)$ as a consequence of the condition that the hypersurface be of constant mean curvature λ (Eisenhart, *loc. cit.*, p. 113). Substituting $r_{ij} = a_i a_j$ in these latter equations we obtain the above equations of Cartan from which it follows that $a_j = \cdots = a_n = a$. Hence from

$$\Sigma_j (b_{ij} - a g_{ij})\lambda_k{}^j = 0 \text{ we have } b_{ij} = a g_{ij} \text{ at } P.$$

Then from the Gauss equations for the hypersurface

$$B_{\alpha\beta\gamma\delta} = b_{\alpha\delta} b_{\beta\gamma} - b_{\alpha\gamma} b_{\beta\delta}$$

we have

$$B_{\alpha\beta\gamma\delta} = a^2 (g_{\alpha\delta} g_{\beta\gamma} - g_{\alpha\gamma} g_{\beta\delta}) \text{ at } P.$$

These equations mean that the hypersurface is of constant curvature provided that the factor a^2 is a constant, i. e. independent of the point P. But multiplying the equations through by $g^{\alpha\delta} g^{\beta\gamma}$ and summing on repeated indices we obtain $\lambda = a^2(n-1)$ from which it follows that a^2 is a constant as required.

PRINCETON UNIVERSITY.

ALMOST PERIODIC FUNCTIONS AND HILL'S THEORY OF LUNAR PERIGEE.*

By AUREL WINTNER.

From the analytical point of view, Hill's theory of the mean motion of the perigee depends on three problems:

(i) determination of the intermediary periodic path of the Moon;

(ii) determination of the characteristic exponent of the orbit thus obtained;

(iii) identification of this characteristic exponent with the mean motion of lunar perigee.

Inequalities which are due to the eccentricities are not considered in Hill's theory and have been developed by Brown. The same holds for the parallactic terms, Hill's equations of motion being derived by a modification of the restricted problem of three bodies.

Hill [1] solved both problems (i), (ii) by a bold application of infinitely many variables, (ii) leading to a certain linear Eigenwertproblem of specific type, and (i) to a non-linear and non-recursive system of infinitely many equations for infinitely many Fourier constants. Hill's application of infinite determinants in the linear problem (ii) has been justified by Poincaré.[2] The question of existence and convergence in case of the non-linear system arising in problem (i) is of quite another nature. Hill was well aware of the novelty of this analytical question and he formulated its mathematical treatment as a desideratum.[3] The proofs requested by Hill were supplied some years ago.[4]

* Received June 20, 1937.

[1] G. W. Hill, *Collected Mathematical Works*, vol. 1 (1905), pp. 284-335 and pp. 243-270. Cf. also the presentation of Poincaré in vol. 2, part 2 (1909), chap. XXV-XXVII of his *Leçons de Mécanique Céleste*.

[2] Cf., e. g., H. Poincaré, *loc. cit.* [1], pp. 49-56.

[3] G. W. Hill, *loc. cit.* [1], p. 287. Poincaré's presentation (*loc. cit.* [1], pp. 31-40) of (i) does not satisfy Hill's desideratum; the legalization of Hill's method in (i) is of a later date.[4] It can be mentioned that the question of the existence of a family of simple closed orbits about the Earth is not identical with the question whether Hill's method of infinitely many variables in the non-linear problem (i) (i. e., the method actually applied in lunar theory) is or is not a procedure which has mathematical legality.

Thus the mathematical problems arising in (i) and (ii) are solved. The present note deals with the mathematical problem in (iii) which presupposes (i) and (ii) in the same way as (ii) presupposes (i).

Hill's justification of the step (iii) is not of a mathematical nature, since it merely consists [5] of an interpretation of Hill's results in terms of the formal theory of Delaunay. Since the Hill-Brown theory is, in the main, independent of the older lunar theories, it seems desirable to give a direct account of the step (iii). From the mathematical point of view, the objection to the usual explanation [5] of (iii) is not merely a matter of method. In fact, it is a well-defined question, whether or not there exists at all a mean motion, and if it does, whether or not its remainder term shows a recurrent behavior which, as tacitly assumed in lunar theory, can be analyzed into an anharmonic Fourier series.

In what follows, these questions will be answered in the affirmative. The proof depends on an appropriate application of a theorem on almost periodic functions which was formulated as a conjecture by the present author and subsequently proved by Bohr.[6] This theorem of Bohr states that if $\zeta(t)$ is a complex-valued almost periodic function of the real variable t and $|\zeta(t)| > \alpha$ holds for some constant $\alpha > 0$ and for every t, then the real continuous function $\arg \zeta(t)$ can be decomposed into the sum of a linear "secular" term μt and of a "recurrent" term, the latter term being almost periodic, while μ is a constant. Here and later on, almost periodicity is meant in the original sense of Bohr.

The problem concerns certain of the solutions of the three Jacobi equations for isoenergetic variations, this system of differential equations being represented by the equation of second order for normal displacements and by the equation of first order for tangential displacements.[7] While the latter

---

According to G. D. Birkhoff ["The Restricted Problem of Three Bodies," *Rendiconti del Circolo Matematico di Palermo*, vol. 39 (1915), part 3, art. 12], the existence of the simple closed orbits in question can be proved in a rather easy way. While such an existence proof was claimed by Poincaré (*loc. cit.*[1], p. 30), it seems to the writer that the passage of the *Méthodes Nouvelles* which is referred to by Poincaré does not imply such a proof, a proof being essentially dependent on Birkhoff's λ-device (cf. H. Poincaré, *loc. cit.*[1], p. 30, line 6) or on some equivalent substitution (facilitating expansions which are *uniformly* convergent in the vicinity of the mass singularity).

[4] A. Wintner, "Zur Hillschen Theorie der Variation des Mondes," *Mathematische Zeitschrift*, vol. 24 (1925), pp. 259-265; also "Ueber die Konvergenzfragen der Mondtheorie," *ibid.*, vol. 30 (1929), pp. 211-227.

[5] G. W. Hill, *loc. cit.*[1], pp. 269-270; also H. Poincaré, *loc. cit.*[1], p. 67.

[6] H. Bohr, "Kleinere Beiträge zur Theorie der fastperiodischen Funktionen, I," *Det Kgl. Danske Videnskabernes Selskab Math.-fys. Meddelelser*, vol. 10, no. 10 (1930).

[7] Cf. G. D. Birkhoff, *loc. cit.*[3], part 1, art. 3.

equation is not needed for the determination of the characteristic exponents, the geometrical problem of the mean motion of lunar perigee cannot be treated by considering the normal displacements alone, instead of the whole problem in the (x, y)-plane. Correspondingly, the situation is different from the one arising in the Adams theory of the mean motion of the lunar node, a problem which is defined in terms of a linear canonical periodic system with a *single* degree of freedom and is, therefore, reducible to the torus problem of Poincaré-Denjoy.[8] In other words,[9] the result to be obtained can be applied to, but is not identical with, an anharmonic Fourier analysis of a sequence of conjugate points on the closed extremal of problem (i), the rotation number being represented by the coefficient μ of the almost periodic remainder term $\arg \zeta(t) - \mu t$.

Hill's equations of motion are of the form

$$x'' - 2y' = \Omega_x(x, y), \qquad y'' + 2x' = \Omega_y(x, y), \tag{1}$$

where the primes indicate derivatives with respect to t. Placing

$$t = m\tau, \tag{2}$$

where m is Hill's parameter determining a solution of (1) within his family of periodic orbits, these periodic solutions are of the form [1]

$$x = \sum_{k=-\infty}^{+\infty} a_k \cos(2k+1)\tau, \quad y = \sum_{k=-\infty}^{+\infty} a_k \sin(2k+1)\tau, \quad \text{where} \quad a_k = a_k(m). \tag{3}$$

The Fourier constants $a_k(m)$ are analytic functions of m and vanish at $m = 0$ with increasing rapidity as $k \to \pm \infty$. Hill's procedure (i) leading to (3) has been justified [4] for $| m | < 0.08333 \cdots$, and so, in particular, for the case $m = +0.08084 \cdots$ of the Moon.

In problem (ii), the validity of Hill's method does not depend on the magnitude of m, since this problem is linear. Exceptional is only the case of an m for which the path (3) has a point on its curve of zero velocity (or reaches the mass singularity). In fact, for this case the equation of normal displacement [10] acquires a singularity, and no direct procedure seems to have

---

[8] T. Levi-Civita, "Sur les équations linéaires à coefficients périodiques et sur le moyen mouvement du noeud lunaire," *Annales de l'Ecole Normale Supérieure*, ser. 3, vol. 28 (1911), pp. 325-376.

[9] Cf. G. D. Birkhoff, "Dynamical systems with two degrees of freedom," *Transactions of the American Mathematical Society*, vol. 18 (1917), part 2, art. 14, and the investigations of Poincaré referred to there.

[10] As to this equation, cf. A. Wintner, "Ueber die Jacobische Differentialgleichung des restringierten Dreikörperproblems," *Berichte der Sächsischen Akademie der Wissenschaften*, vol. 82 (1930), pp. 345-354.

been developed for the resulting singular Eigenwertproblem. The first positive [11] singular value of m is, however, quite large and is represented by the value $m = 0.56 \cdots$ belonging to Hill's cuspidal orbit of maximum lunation,[12] a value beyond the m-range under consideration.

The Jacobi equations belonging to (1) and (3) can be written as

$$\begin{aligned} \xi'' - 2\eta' &= \Omega_{xx}(t; m)\xi + \Omega_{xy}(t; m)\eta, \\ \eta'' + 2\xi' &= \Omega_{xy}(t; m)\xi + \Omega_{yy}(t; m)\eta, \end{aligned} \tag{4}$$

and have, for fixed m, four linearly independent solutions $(\xi(t), \eta(t))$ two of which are of a trivial type, while the remaining two, which represent isoenergetic displacements, supply the general solution of the equation of (isoenergetic) normal displacements; cf. (ii). The pair of trivial solutions of (4) is obtained by differentiating the solution (3) of (1), where $\tau = t/m$, with respect to m and the origin of the t-axis respectively, so that these solutions of (4) belong to the trivial pair $(0, 0)$ of characteristic exponents and the first of these two solutions contains a secular term and is not an isoenergetic displacement. Since the orbit (3) is symmetrical with respect to both coördinate axes $y = 0$, $x = 0$ and is formally stable in the m-domain under consideration, it is readily inferred from the Fuchs-Floquet theory that, if $(c, -c)$ represents the pair of non-trivial characteristic exponents, the two non-trivial solutions of (4) can be represented in the form

$$\begin{aligned} \xi(t; m) &= \gamma \sum_{k=-\infty}^{+\infty} A_k \cos\{(2k + 1 + c)\tau + \delta\}, \\ \eta(t; m) &= \gamma \sum_{k=-\infty}^{+\infty} B_k \sin\{(2k + 1 + c)\tau + \delta\}, \end{aligned} \tag{5}$$

---

[11] As to the corresponding range of the family (3) in the retrograde case $m < 0$, cf. G. D. Birkhoff, *loc. cit.* [3], part 4, art. 16.

[12] While Hill [1] had to derive this orbit, not from his analytical expansions, but by using mechanical quadratures, a mathematical existence proof for this orbit can be considered as implied by Strömgren's general termination principle, the existence of the family (3) being established for small $m > 0$. Unfortunately, sufficient material is not available concerning the natural termination of Hill's family, and the same holds precisely for that among the Copenhagen groups which corresponds to Hill's family, i. e., for Strömgren's Klasse g. The analytic continuation of the lunar orbits (3) which is given by Kelvin in vol. 4, no. 55, of his *Collected Works*, is rather flat with fully developed loops (the latter arising at the stage of Hill's cuspidal orbit). Judging from Kelvin's single orbit, which may or may not be sufficiently advanced, one possibility for the natural termination would be an indefinite flattening in the direction perpendicular to the axis of syzygies, the orbit becoming the whole y-axis, while the Sun is infinitely distant and on the x-axis.

where $\gamma \neq 0$ and δ are arbitrary real integration constants and

$$c = c(m); \qquad A_k = A_k(m), \qquad B_k = B_k(m) \qquad (m = t/\tau)$$

are real analytic functions of m.

As $m \to 0$, the behavior of $A_k(m)$ and $B_k(m)$ in (5) is, for large $|k|$, similar to that of $a_k(m)$ in (3). In particular, the limiting values $(A_k)_{m=0}$ and $(B_k)_{m=0}$ vanish unless $|2k+1| = 1$. If $|2k+1| = 1$, so that $k = 0$ or $k = -1$, it is found from the expansions of $A_k(m)$ and $B_k(m)$ that

$$(6_1) \qquad A_0^0 \neq 0, \quad B_0^0 = A_0^0; \qquad A_{-1}^0 = -3A_0^0, \quad B_{-1}^0 = 3B_0^0,$$

where $A_k^0 = (A_k)_{m=0}$, $B_k^0 = (B_k)_{m=0}$. The limit relations (6_1), together with

$$(6_2) \qquad A_k^0 = 0, \quad B_k^0 = 0, \quad \text{where} \quad 2k+1 \neq \pm 1,$$

have been checked by comparison with Kepler's motion.

Letting $m \to 0$ in (5), using (6_1)–(6_2) and omitting the multiplicative integration constant $\gamma \neq 0$ (or, rather, $2A_0^0\gamma \neq 0$), one finds after an easy reduction that, if c^0 denotes the limiting value $(c)_{m=0}$ of the characteristic exponent $c = c(m)$,

$$(\xi)_{m=0} = -\cos\tau\cos(c^0\tau+\delta) - 2\sin\tau\sin(c^0\tau+\delta),$$
$$(\eta)_{m=0} = -\sin\tau\cos(c^0\tau+\delta) + 2\cos\tau\sin(c^0\tau+\delta).$$

Hence

$$(\xi^2+\eta^2)_{m=0} = \cos^2(c^0\tau+\delta) + 4\sin^2(c^0\tau+\delta).$$

Since this continuous and non-vanishing function of τ is periodic, there exists a sufficiently small constant α_0 which depends on the integration constants δ and $\gamma \neq 0$ but is such that

$$(\xi^2+\eta^2)_{m=0} > \alpha_0 > 0 \quad \text{for} \quad -\infty < \tau < +\infty.$$

Hence there exists a sufficiently small $m^* > 0$ such that if m has any value between $-m^*$ and m^*, one can choose a positive constant $\alpha = \alpha_m$ for which the almost periodic functions (5) of τ or $t = m\tau$ satisfy the inequality

$$(7) \qquad [\xi(t;m)]^2 + [\eta(t;m)]^2 > \alpha_m > 0; \qquad -\infty < t < +\infty,$$

it being understood that α_m is, for fixed m, a function of the integration constants γ, δ occurring in (5), while m^* is independent of γ, δ.

As pointed out above, the coefficients $A_k(m)$ and $B_k(m)$ tend, for fixed

small values of m, rather strongly to zero as $k \to \pm \infty$. Correspondingly, it requires only quite rough numeral estimates to show that (7) is satisfied for the value of m which belongs to the Moon, i. e., that $m = 0.08084 \cdots$ is "sufficiently small" or less than m^*.

For a fixed m and for fixed values of the integration constants $\gamma \neq 0$, δ, introduce the perihelium ϖ, or rather perigee, as a function of t or τ by placing

$$\xi = \rho \cos \varpi, \quad \eta = \rho \sin \varpi, \quad \text{where} \quad \rho = (\xi^2 + \eta^2)^{\frac{1}{2}} \geqq 0. \tag{8}$$

It is understood that, since $\rho \neq 0$ by (7), the value of $\varpi = \varpi(t)$ is determined for every t by continuity, if one assigns an initial normalizing condition, say $0 \leqq \varpi(0) < 2\pi$. Put

$$\zeta(t) \equiv \zeta = \xi + i\eta = (\xi^2 + \eta^2)^{\frac{1}{2}} \exp i\varpi; \text{ cf. (8).} \tag{9}$$

Then $\zeta(t)$ is, in view of (5) and (2), an almost periodic function. Furthermore, (7) shows that $|\zeta(t)| > \alpha_m^{\frac{1}{2}} > 0$ for every t. Consequently, the theorem of Bohr [6] is applicable to $\zeta(t) = |\zeta| \exp i\varpi$. It follows that the perigee $\varpi = \varpi(t)$ is of the form

$$\varpi(t) = \mu t + \text{an almost periodic function of } t, \tag{10}$$

where μ is a constant.

It is instructive to look upon (10) from the point of view of the fundamental mathematical difficulty of celestial mechanics, a difficulty represented by the problem of small integration divisors in the astronomical theories and by Birkhoff's unsolved stability problem of incommensurable rotation numbers in the corresponding mathematical theories. First, from (8),

$$\varpi' \equiv \frac{d\varpi}{dt} = \frac{\eta'\xi - \xi'\eta}{\xi^2 + \eta^2}. \tag{11}$$

It follows, therefore, from (7), (5) and (2) that the derivative $\varpi'(t) \equiv \varpi'(m\tau)$ of the perigee is an almost periodic function. Now the integral of an almost periodic function is almost periodic whenever it is bounded. Hence, if μ denotes the constant term in the Fourier series of the almost periodic function (11), so that

$$\mu = \lim_{T \to \infty} \frac{1}{T} \int_0^T \varpi'(t)\, dt, \tag{12}$$

then (10) can be inferred from (11) by integration if and only if $\varpi(t) - \mu t$ is a bounded function. On the other hand, the integral of an almost periodic function which has no constant term in its Fourier series is, according to Favard, sure to be almost periodic if there are no frequencies of arbitrarily small absolute value, i. e., if there are no "small divisors"; while otherwise one cannot say anything, as shown precisely by the classical astronomical problem. Now all that follows from (11), (7), (5) and (2) is that the frequencies of the almost periodic function $\varpi'(m\tau)$ or $\varpi'(m\tau) - \mu$ are contained in the double sequence

$$\nu_1 c + \nu_2, \quad \text{where} \quad \nu_1, \nu_2 = 0, \pm 1, \pm 2, \cdots. \tag{13}$$

Since $c = c(m)$ is a continuous and non-constant function of m and can, therefore, be considered as irrational, the double sequence (13) is identical with the sequence of frequencies occurring in the classical problem of small divisors. Accordingly, Bohr's theorem on $\arg \zeta(t)$ plays, in the above proof of (10), a definite dynamical rôle: it assures that the classical small divisors are harmless in the present case [although the continuous function $c = c(m)$ in (13) attains values which are represented by arbitrarily irregular continued fractions, no c-set of measure zero being excluded]. This could be checked by a direct discussion of the Fourier constants of the almost periodic function (11).

Since all frequencies of the almost periodic function $\varpi'(m\tau)$ of $\tau = t/m$ are contained in (13), the Fourier series of the almost periodic function on the right of (10) is of the form

$$\varpi(t) - \mu t \sim \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} C_{jk} \cos \left\{ \frac{c(m)j + k}{m} t + \delta_{jk} \right\}. \tag{14}$$

This agrees with a result of Bohr,[13] according to which both the frequencies and the secular constant of $\varpi(m\tau)$ must be of the form (13); so that

$$\mu = (cr + s)/m, \qquad c = c(m), \tag{15}$$

where r and s are integers.

Hill did not consider the Fourier series (14) but rather the constant μ of the secular part of the perigee $\varpi(t)$; and he wrote,[14] corrrespondingly,

---

[13] H. Bohr, "Ueber fastperiodische ebene Bewegungen," *Commentarii Mathematici Helvetici*, vol. 4 (1932), pp. 51-64.

[14] G. W. Hill, *loc. cit.* [1], pp. 269-270. Hill's notations are not the same as those used above; the difference is explained below.

$\varpi(t) = \mu t$ and $\varpi'(t) = \mu$. It is, however, obvious from the context that what Hill actually had in mind were relations of the type (14), (10) and (12). While this assumption of Hill is now justified, the corresponding mathematical postulates in Brown's theory, where the higher powers of the lunar eccentricity are not neglected, seem to require an entirely different approach.[15]

While (8) refers the perigee $\varpi(t)$ to the synodical coördinate system, Hill's perigee $\omega(t)$ is referred to a siderial coördinate system. Since the first of these frames, when considered from the second, rotates with constant angular velocity, $\varpi(t) - \omega(t)$ is a linear function of t, and so the above results hold for $\omega(t)$ also. Since $c = c(m)$ and $\mu = \mu(m)$ are continuous functions of m, while r and s in (15) are integers, one can determine the actual values of r, s for small m by letting $m \to 0$ and comparing the result with the formulae of Kepler's motion. This, when combined with the reduction of $\varpi(t)$ to $\omega(t)$, gives as connection between the mean motion of lunar perigee and the characteristic exponent c precisely the relation stated by Hill.[5]

THE JOHNS HOPKINS UNIVERSITY.

[15] Cf. A. Wintner, *loc. cit.* [4], pp. 214-215.

SUR LES MOUVEMENTS DES SYSTÈMES DYNAMIQUES QUI ADMETTENT " L'INCOMPRESSIBILITÉ " DES DOMAINES.*

Par HEINRICH HILMY.

1. Considérons un système dynamique dont les mouvements sont définis par l'équation différentielle

$$(1) \qquad dx_i/dt = X_i(x_1, \cdots, x_n) \qquad (i = 1, \cdots, n).$$

Nous interprétons la variable indépendante t comme le temps et les variables $x_1, \cdots, x_n$, comme les coordonnées du point p se mouvant dans l'espace euclidien E^n, à n-dimensions.

Supposons que les seconds membres des équations (1) sont des fonctions continues et qu'ils satisfont aux conditions de Lipschitz dans les points des sous-ensembles considérés E^n. Dans ce cas auront lieu les théorèmes généraux de l'existence et de l'unicité des solutions, et également le théorème de la dépendance continue des solutions des conditions initiales.

Les mouvements du système dynamique sont caractérisés par les trajectoires que décrivent dans E^n les points $p \subset E^n$. Nous désignons par $f(p, t)$ la position du point p dans le moment du temps t; $f(p, 0) = p$.

Si A est un ensemble contenu dans E^n le symbole $f(A, t)$ désigne l'ensemble $\{f(p, t)\}$ où $p \subset A$.

Le symbole $S(x, \epsilon)$ désigne le voisinage sphérique du point p au rayon ϵ.

L'ensemble M est dit positivement (négativement) invariant, si $f(p, t) \subset M$ pour toute valeur de $t \geq 0$ $(t \leq 0)$. L'ensemble M est dit invariant s'il est en même temps positivement et négativement invariant.

Le mouvement $f(p, t)$ est dit positivement (négativement) stable au sens de Lagrange si l'on peut indiquer une sphère S dont le centre serait le point p et le rayon sufisemment grand pour que $f(p, t) \subset S$ pour tout $t \geq 0$ $(t \leq 0)$.

Le mouvement $f(p, t)$ est dit positivement (négativement) stable au sens de Poisson si pour tous arbitraires $\epsilon > 0$ et $T > 0$ on peut indiquer des valeurs de t positives (négatives) qui satisfassent à l'inégalité $| t | > T$ et pour lesquels $\rho(p, f(p, t)) < \epsilon$ [$\rho(x, y)$ désigne la distance entre les points x et y dans E^n]. Le mouvement $f(p, t)$ est dit stable au sens de Poisson s'il est tel simultanement dans le sens positive et dans le sens négative.

Convenons de dire que le mouvement instable au sens de Poisson et en

* Received May 3, 1937.

même temps stable au sens de Lagrange est un mouvement asymptotique. Si $f(p, t)$ est un mouvement positivement (négativement) instable au sens de Poisson, il découle de la continuité des seconds membres de l'équation (1) que pour tout point p' de la trajectoire $f(p, t)$ on peut indiquer un voisinage sphérique suffisemment petit pour que le point p', ayant une fois quitté ce voisinage n'y revienne jamais le temps changeant dans le sens positif (négatif). Nous dirons que le nombre positif η égal à la borne supérieure des rayons de tous les voisinages sphériques du point p' dans lesquels il ne revient plus après l'avoir une fois quitté, le temps changeant dans le sens positif (négatif),— nous dirons que ce nombre est l'indice caractéristique du point p' pour le temps croissant (décroissant).

2. Nous dirons que dans l'ensemble invariant M a lieu l'incompressibilité des domaines si pour tout domaine relatif borné $G \subset M$ se vérifie l'énoncé suivant: pour aucune valeur de t le domaine $f(G, t)$ n'est contenu dans une partie vraie de G.

THÉORÈME. *Si dans l'ensemble fermé invariant M a lieu " l'incompressibilité " des domaines, tous les points de cet ensemble, sauf l'ensemble des points de la première catégorie de R. Baire, appartiennent à la somme des ensembles des points situés sur les trajectoires de mouvements stables au sens de Poisson et de ceux des points situés sur les trajectoires des mouvements instables au sens de Lagrange dans les deux sens.*

Démontrons, pour commencer, le théorème pour le temps croissant.

Désignons par Q l'ensemble de tels points $p \subset M$ qui sont situés sur les trajectoires des mouvements asymptotiques pour $t \to +\infty$. Prenons une suite de nombres positifs croissant indéfiniment et uniformement

$$R_1, R_2, \cdots, R_n, \cdots$$

et désignons par

$$S_1, S_2, \cdots, S_n, \cdots$$

l'ensemble des points $p \subset M$ situés dans les sphères et sur leurs frontières aux rayons respectivement égaux à $R_1, R_2, \cdots, R_n, \cdots$ décrites autour du point arbitraire appartenant à l'ensemble M.

Soit Q^k un ensemble de tels points $p \subset Q \cdot S_k$ pour lesquels on a $f(p, t) \subset S_k$ pour toute valeur de $t \geq 0$; alors

$$Q = \sum_{i=1}^{\infty} Q^i. \tag{2}$$

Montrons que chacun des ensembles Q^k est un ensemble de la première catégorie de Baire par rapport à S_k et par conséquent par rapport à M également.

Donnons nous une suite de nombres positifs

$$\epsilon_1, \epsilon_2, \cdots, \epsilon_n, \cdots$$

convergeant uniformément à zéro. Désignons par Q_m^k l'ensemble des points $p \subset Q^k$ dont l'indice caractéristique est $\eta \geq \epsilon_m$. Il est évident que

$$Q^k = \sum_{i=1}^{\infty} Q_i^k. \tag{3}$$

Démontrons que pour k fixé et pour tout m l'ensemble Q_m^k n'est dense nulle part dans S_k. Admettons le contraire; il se trouverait alors un domaine relatif $G \subset S_k$, dans lequel l'ensemble de points appartenant à Q_m^k serait partout dense.

Donnons nous un nombre arbitraire $\tau > 0$; choisissons dans le domaine G le point $p' \subset Q_m^k$ et décrivons autour de ce point la sphère S^* au rayon $\epsilon < \frac{1}{2}\epsilon_m$ et suffisemment petit pour que les conditions $S^* \subset G$ et $f(S^*, \tau) \cdot S^* = 0$ soient vérifiées, ce qui est possible en vertu de la dependance continuelle des conditions initiales. Montrons à present que $f(S^*, t) \cdot S^* = 0$ pour tout $t \geq \tau$. S'il n'était pas ainsi, on pourrait choisir un point $q \subset S^*$ et un nombre $T > \tau$ tels que $f(q, T) \subset S^*$. Soit $\epsilon^* > 0$ suffisemment petit pour que $S(q, \epsilon^*) \subset S^*$. En vertu de la dépendance continue des trajectoires des conditions initiales on peut indiquer un nombre $\delta > 0$ suffisemment petit pour que

$$f(S(q, \delta), T) \subset S(f(q, T), \epsilon^*).$$

Mais $S(q, \delta)$ contient nécessairement le point $p^* \subset Q_m^k$, $f(p^*, \tau) \not\subset S^*$ et $f(p^*, T) \subset S^*$ par conséquent le point p^*, ayant quitté la sphère S^* y revienne encore, ce qui est contradictoire vu que l'indice caractéristique de p^* surpasse ou égale ϵ_m.

Considérons le domaine

$$\Gamma = \sum_{t \geq 0} f(S^*, t).$$

Le domaine Γ est positivement invariant et borné,

$$f(\Gamma, \tau) = \sum_{t \geq \tau} f(S^*, t)$$

étant un sous-ensemble de Γ vrai vu que $S^* \subset \Gamma$ et $f(\Gamma, \tau) \cdot S^* = 0$ ce qui contredit la supposition que " l'incompressibilité " des domaines a lieu dans M. La contradiction à laquelle nous arrivons montre que l'ensemble Q_m^k n'est

dense nulle part dans S_k. Mais alors l'ensemble Q^k, en vertu de (3) et l'ensemble Q en vertu de (3) seront des ensembles de la première catégorie de Baire par rapport à M.

En désignant par N l'ensemble des points $p \subset M$ qui sont situés sur des trajectoires asymptotiques pour $t \to -\infty$ et en répétant les considérations faites plus haut, démontrons que l'ensemble N sera un ensemble de la première catégorie de R. Baire par rapport à M. Mais alors l'ensemble $A = Q + N$ sera également un ensemble de la première catégorie de R. Baire par rapport à M, et le théorèm est démontré.

3. Le théorème démontré complète les résultats obtenus dans le mémoire connu de E. Hopf [1] qui traite de la généralisation du Wiederkehrsatz de H. Poincaré en élucidant la catégorie au sens de R. Baire des ensembles étudiés dans ce mémoire. Le théorème de E. Hopf est applicable aux systèmes dynamiques qui admettent l'invariant intégral de H. Poincaré. Le théorème que nous venons de démontrer est applicable à une classe plus large de systèmes dynamiques, vu que l'incompressibilité des domaines peut avoir lieu dans des systèmes dynamiques qui n'admettent pas d'invariant intégral. Construisons un exemple d'un système dynamique qui justifie cette assertion.

Considérons un segment de l'axe de x, de la longueur 1 aux coordonnées des extrémités $-\frac{1}{2}$ et $+\frac{1}{2}$ et construisons sur ce segment un ensemble formé et partout non dense que nous désignons par P.

Prenons une suite de nombres

$$\alpha_n = 1/4^n \qquad (n = 1, 2, \cdots).$$

Construisons sur le premier segment un autre segment de la longueur α_1, de manière que son centre coïncide avec le centre du premier segment et excluons tous les points intérieurs du second segment. Sur chacun des deux segments conservés construisons de nouveau des segments de la longueur α_2, de manière que les centres de ces segments coïncident avec les centres des segments qui les contiennent; excluons ensuite encore tous les points intérieurs des segments que nous venons de construire. Représentons nous ce procédé indéfini. Désignons par G l'ensemble de tous les points du segment $[-\frac{1}{2}, +\frac{1}{2}]$ qui peuvent être exclus par ce procédé et par P l'ensemble de tous les autres points.

Il est évident que l'ensemble P est un ensemble fermé et partout non dense, tandis que l'ensemble G est dense partout sur le premier segment.

De simples calculs nous montrent que

[1] E. Hopf, "Zwei Sätze über den wahrscheinlichen Verlauf der Bewegungen dynamischer Systeme," *Mathematische Annalen*, Bd. 103 (1930).

$$\text{mes } P = \tfrac{1}{2}, \qquad \text{mes } G = \tfrac{1}{2}.$$

L'ensemble des intervalles, exclus du segment $[-\frac{1}{2}, +\frac{1}{2}]$ lors de la construction de l'ensemble P est un ensemble dénombrable. Donc nous pouvons les énumérer. Soit

$$\Delta_1, \Delta_2, \cdots, \Delta_n, \cdots$$

la suite de ces intervalles. Désignons par a_n et par b_n les coordonnées respectives de l'extrémité droite et de l'extremité gauche de l'intervalle Δ_n. Définissons sur le segment $[-\frac{1}{2}, +\frac{1}{2}]$ la fonction $\Phi(x)$ comme il suit:

$$\Phi(x) = \begin{cases} 0 \quad , & \text{si } x \subset P, \\ x - a_n, & \text{si } a_n \leq x \leq a_n + \dfrac{b_n - a_n}{2}, \\ b_n - x, & \text{si } a_n + \dfrac{b_n - a_n}{2} \leq x \leq b_n. \end{cases}$$

La fonction $\Phi(x)$ est une fonction continue sur tout le segment $[-\frac{1}{2}, +\frac{1}{2}]$. De plus on peut constater par des simples calculs que pour deux points quelconques x_1 et x_2 du segment $[-\frac{1}{2}, +\frac{1}{2}]$ se vérifie l'inégalité:

$$|\Phi(x_1) - \Phi(x_2)| \leq |x_1 - x_2|$$

c'est à dire que la fonction $\Phi(x)$ satisfait à la condition de Lipschitz.

Examinons dans le plan euclidean E^2 l'ensemble de points M:

$$-\tfrac{1}{2} \leq x \leq +\tfrac{1}{2},$$
$$-\infty \leq y \leq +\infty.$$

Construisons sur le segment $[-\frac{1}{2}, +\frac{1}{2}]$ l'ensemble P et définissons ensuite sur ce segment la fonction $\Phi(x)$.

Définissons sur l'ensemble M un système dynamique, dont le mouvement est exprimé par l'équation différentielle suivante:

$$dx/dt = 0, \qquad dy/dt = -[|y| + \Phi(x)].$$

Les seconds membres de ces équations différentielles sont des fonctions continues satisfaissant aux conditions de Lipschitz.

Les lignes sur lesquelles sont situées les trajectoires des points de l'ensemble M sont les droites:

$$x = \text{const.}$$

Les trajectoires des mouvements sont les ensembles suivants: tous les points appartenant à l'ensemble P construit sur l'axe x sont des points de repos,

vu que dans ces points les seconds membres des équations différentielles s'annulent. Les droites qui n'ont pas de parties communes avec l'ensemble P sont des trajectoires de mouvements instables au sens de Lagrange dans les deux sens. L'ensemble de points situés sur ces droites est un ensemble dense partout dans M. Les droites qui ont des parties communes avec l'ensemble P sont formées par les trajectoires de trois mouvements indépendants: du point de repos sur l'axe x et de deux semi-droites qui y sont contigues et qui sont les trajectoires de deux mouvements asymptotiques. Le point de repos est pour l'un de ces mouvements le point limite alpha, et pour l'autre—le point limite oméga du mouvement.

Désignons par M^* un ensemble de points dont les coordonnées satisfont aux inégalités $-\frac{1}{2} \leq x \leq +\frac{1}{2}$, $0 \leq y \leq 1$ et qui de plus sont situés sur des droites ayant des parties communes avec l'ensemble P. Les ponits de l'ensemble M^*, pour lesquels $y = 0$ sont des points de repos, et les points pour lesquels $y > 0$ se rapprochent asymptotiquement des points de repos pour $t \to +\infty$.

L'ensemble M^* est mesurable: mes $M^* = \frac{1}{2}$. Il découle du caractère des mouvements des points de l'ensemble M^* que

$$\lim_{t \to +\infty} \text{mes}\, f(M^*, t) = 0.$$

Ceci veut dire que l'invariant intégral de H. Poincaré ne peut être introduit dans le système dynamique éxaminé. Mais dans ce système a lieu l'incompressibilité des domaines, vu qu'il se trouvera dans chacun de ces domaines un point qui quitte ce domaine pour toujours.

Moscou, U. S. S. R.

THE STRUCTURE OF MONOTONE FUNCTIONS.*

By PHILIP HARTMAN and RICHARD KERSHNER.

Introduction. The present paper deals with continuous, monotone increasing functions

$$y = f(x), \text{ where } 0 \leqq x \leqq 1 \text{ and } 0 \leqq y \leqq 1,$$

and attempts a classification of the absolutely continuous or singular behavior of these functions, these notions being meant in the sense of Lebesgue. It is clear that this class of functions is in one-to-one correspondence with the class of monotone mappings of a dense enumerable subset of the interval $0 \leqq x \leqq 1$ into a dense enumerable set of the interval $0 \leqq y \leqq 1$. Hence it is natural to attempt to describe the Lebesgue properties (absolute continuity, pure "singularity," or mixed behavior) of these functions $f(x)$ directly in terms of the asymptotic or qualitative properties of the two dense sequences of numbers which are mapped on to each other by $y = f(x)$. This approach was suggested to us by Professor Wintner.

One can fix one of the two sequences arbitrarily without loss of generality. The simplest choice seems to be the dense set on the x-interval consisting of all points which have a finite dyadic development. Thus the Lebesgue properties of $f(x)$ will be explicitly described in terms of the sequence of numbers $\{y_j\}$ consisting of all function values $f(k/2^n)$, where $k = 0, 1, \cdots, 2^n$ and $n = 0, 1, \cdots$. For instance, there will be obtained (Section 1) necessary and sufficient conditions for the function $y = f(x)$ to be absolutely continuous and corresponding conditions for $y = f(x)$ to be purely singular, i. e., to have no absolutely continuous component.

Essential use will be made of the fact that, while the inverse of a strictly increasing, absolutely continuous function need not be absolutely continuous, the inverse of a strictly increasing, purely singular function is necessarily purely singular.

For a certain class of functions, which will be called the class belonging to the symmetric case, the results will be particularly complete (Sections 3 and 6). For instance, both these functions and their inverses will be shown to be "pure," i. e., either absolutely continuous or purely singular. In the treatment of the symmetric case, use is made of the "pure" theorem [1] of the

---

* Received April 23, 1937.

[1] Jessen and Wintner [4], Theorem XXXV.

theory of infinite convolutions and the theory of probable convergence of series of random variables. Actually, infinite products of random variables are used, but this is easily reduced to the theory of series. The latter theory, developed by Rademacher [12], Khintchine and Kolmogoroff [6] and others, is known to be equivalent to the convergence theory of infinite convolutions.[2] The theory of convolutions is made applicable by showing (Section 6) that certain of the functions of the symmetric case are infinite convolutions of the Poisson type; cf. Wintner [15]. This fact, on the one hand, makes possible a sharpening of the results of Section 3 and, on the other hand, provides an interesting theorem on infinite convolutions.

The description of the function $y = f(x)$ in terms of the sequence of numbers $k/2^n$ provides for the simple construction of particular functions of a given type, for instance, (Section 4) absolutely continuous functions which are nowhere monotone, or (Section 5) strictly increasing, purely singular functions which show that no restriction on the modulus of continuity which is weaker than a uniform Lipschitz condition of order 1 assures absolute continuity.

The theorems to be proved imply, in particular, the results and examples given by Brodén [1], Minkowski [10], Denjoy [2], Faber [3], Sierpiński [13], van Kampen and Wintner [5]. The products $\Pi(1 \pm \epsilon_n)$ occurring in the sequel have been considered by Brodén [1] and their connection with the question of singularity has been recognized by Faber [3] and more particularly by Rademacher [11].

1. The general case. Let $y = f(x)$ be a continuous, monotone increasing function defined on the interval $[0, 1]$ such that $f(0) = 0$ and $f(1) = 1$. Let y_n^k for $k = 0, 1, \cdots, 2^n$ denote the image of the point $x = k/2^n$ so that

$$f(k/2^n) = y_n^k; \qquad (k = 0, 1, \cdots, 2^n;\ n = 0, 1, \cdots). \tag{1}$$

Then the set $\{y_n^k\}$ is a dense set R of points on the interval $[0, 1]$, and, since $f(x)$ is monotone increasing,

$$y_n^k \begin{matrix} \leqq \\ \geqq \end{matrix} y_m^j \text{ according as } k/2^n \begin{matrix} \leqq \\ \geqq \end{matrix} j/2^m. \tag{2}$$

The numbers y_n^k, where $k = 0, 1, \cdots, 2^n$, will be called the n-th section R_n of R.

Denote by $\gamma_n^0, \gamma_n^1, \cdots, \gamma_n^{2^n-1}$ the 2^n open intervals

$$\gamma_n^k: \quad y_n^k < y < y_n^{k+1}; \qquad (k = 0, 1, \cdots, 2^n - 1), \tag{3}$$

into which the y-interval $[0, 1]$ is divided by the points of R_n. Each interval γ_n^k is subdivided into exactly two subintervals γ_{n+1}^{2k} and γ_{n+1}^{2k+1}. For convenience,

[2] Jessen and Wintner [4], Theorem XXXII.

one of these subintervals will be denoted by α_n^k, the other by β_n^k, the notation being fixed so that

$$\alpha_n^k \geqq \beta_n^k \text{ and } \alpha_n^k + \beta_n^k = \gamma_n^k;\ k = 0, 1, \cdots, 2^n - 1, \tag{4}$$

where the same letter is used to denote either the interval or its length. Each point y of $[0, 1]$ which is not a point of R_n determines, for every n, a unique interval γ_n^k such that y is in γ_n^k. The corresponding sequence of intervals will be denoted by $\gamma_n(y)$. The intervals α_n^k and β_n^k into which the interval $\gamma_n^k = \gamma_n(y)$ is subdivided by a point of R_{n+1} will be denoted by $\alpha_n(y)$ and $\beta_n(y)$. Now let

$$\begin{aligned} \epsilon_n(y) &= [\alpha_n(y) - \beta_n(y)]/\gamma_n(y) \text{ if } y \text{ is in } \alpha_n(y) \\ \epsilon_n(y) &= [\beta_n(y) - \alpha_n(y)]/\gamma_n(y) \text{ if } y \text{ is in } \beta_n(y) \end{aligned} \qquad (n = 0, 1, \cdots). \tag{5}$$

The symbol $\epsilon_n(y)$ will remain undefined if y is a point of R_{n+1}. Thus $\epsilon_n(y)$ is a step function defined over the complement of R_{n+1} in $[0, 1]$; also $-1 < \epsilon_n(y) < +1$.

THEOREM I. *The continuous, strictly increasing function $y = f(x)$ is purely singular, mixed, or absolutely continuous according as the set of points y for which the infinite product*

$$\prod_{n=0}^{\infty} [1 + \epsilon_n(y)] \tag{6}$$

diverges,[3] is of measure one, of measure between one and zero, or of measure zero.

The proof of Theorem I depends on the following

LEMMA I. *A continuous function of bounded variation, $y = \phi(x)$, is purely singular, mixed, or absolutely continuous according as the variation on the set of points at which the derivative $\phi'(x)$ is not finite is equal to the total variation, less than the total variation but positive, or zero.*

This lemma is implicitly contained in the work of Lebesgue [9].

Proof of Theorem I. In order to apply Lemma I, notice that the variation over the set of points at which $f'(x)$ is not finite is equal to the measure of the set of points y for which $df^{-1}(y)/dy$ is zero, where $x = f^{-1}(y)$, $0 \leqq y \leqq 1$, is the continuous, monotone, inverse function of $y = f(x)$, $0 \leqq x \leqq 1$. Put

$$h_n(y) = (\tfrac{1}{2})^n/\gamma_n(y), \qquad (n = 0, 1, \cdots), \tag{7}$$

[3] Divergence of an infinite product is understood to include the case of divergence to zero.

so that, by (1), (3), and the definition of $\gamma_n(y)$,

(8) $h_n(y) = [f^{-1}(y_n^{k+1}) - f^{-1}(y_n^k)]/(y_n^{k+1} - y_n^k)$, where $y_n^k < y < y_n^{k+1}$.

It is clear from (8) that $h_n(y) \to df^{-1}(y)/dy$, $n \to \infty$, if $df^{-1}(y)/dy$ exists. Since $f^{-1}(y)$ is monotone, its derivative exists almost everywhere on the interval $0 \leqq y \leqq 1$. Thus

(9) $$h_n(y)\,(\to) df^{-1}(y)/dy, n \to \infty,$$

where the parenthesis enclosing a relation sign means that the indicated relation holds almost everywhere.

Now suppose that $\gamma_{n+1}(y) = \alpha_n(y)$; then, by (5),

$$\alpha_n(y) - \beta_n(y) = \epsilon_n(y)\gamma_n(y).$$

Hence, from (4),

$$\alpha_n(y) = \tfrac{1}{2}\gamma_n(y)[1 + \epsilon_n(y)].$$

Similarly, if $\gamma_{n+1}(y) = \beta_n(y)$, then

$$\beta_n(y) = \tfrac{1}{2}\gamma_n(y)[1 + \epsilon_n(y)].$$

Thus, in any case,

$$\gamma_{n+1}(y) = \tfrac{1}{2}\gamma_n(y)[1 + \epsilon_n(y)],$$

and so, since $\gamma_0(y) \equiv 1$,

(10) $$\gamma_{n+1}(y) = (\tfrac{1}{2})^{n+1} \prod_{k=0}^{n} [1 + \epsilon_k(y)].$$

From (7), (9) and (10),

$$1/\prod_{k=0}^{n} [1 + \epsilon_k(y)]\,(\to) df^{-1}(y)/dy, \qquad n \to \infty,$$

that is,

(11) $$1/\prod_{k=0}^{\infty} [1 + \epsilon_k(y)]\,(=) df^{-1}(y)/dy.$$

Theorem I now follows from (11) and Lemma I in virtue of the remark made at the beginning of the proof.

2. A criterion for pure singularity. A weakened form of the criterion for pure singularity furnished by Theorem I may be given which lends itself readily to the explicit calculation of particular examples.

THEOREM II. *If* $\liminf_{n\to\infty} \beta_n(y)/\alpha_n(y)\,(<)1$, *then* $y = f(x)$ *is purely singular.*

Proof. Since, by (5),

$$|\epsilon_n(y)| = \left|\frac{\alpha_n(y) - \beta_n(y)}{\alpha_n(y) + \beta_n(y)}\right| > \tfrac{1}{2}\left|1 - \frac{\beta_n(y)}{\alpha_n(y)}\right|,$$

it is clear that $\liminf_{n\to\infty} [\beta_n(y)/\alpha_n(y)](<)1$ implies $\limsup_{n\to\infty} |\epsilon_n(y)|(>)0$, which in turn implies the divergence of the infinite product (6) almost everywhere on the interval $0 \leqq y \leqq 1$. Hence Theorem II follows from Theorem I.

A continuous, monotone increasing function considered by Minkowski [10] in connection with Lagrange's theorem on the continued fractions of quadratic irrationalities was proved to be purely singular by Denjoy [2]. Denjoy's proof depended on the metric properties of continued fractions. A more direct treatment can be obtained from Theorem II as follows:

For convenience, we shall denote by P_n a point of R_n which is not a point of R_{n-1}. The Minkowski function is determined by the following set R: Let R_0 consist of the points $0/1$ and $1/1$; suppose that R_n has been defined and consists of rational numbers, let p/q, p'/q' be two successive points of R_n, then $(p+p')/(q+q')$ is the point P_{n+1} between p/q and p'/q'. Note that if $p/q < p'/q'$ are two successive points of R_n, then $p'q - pq' = 1$, so that the $\gamma_n{}^k$ determined by these two points is of length $1/(qq')$. Thus the ratio $\beta_n{}^k/\alpha_n{}^k$ is q/q' or q'/q according as $q < q'$ or $q' < q$. In particular, the interval γ_{n+2}^{4k} determined by R_{n+2} whose endpoints are p/q and $(2p+p')/(2q+q')$ is divided by a point P_{n+3} in such a way that

$$\beta_{n+2}^{4k}/\alpha_{n+2}^{4k} = q/(2q+q') < \tfrac{1}{2}.$$

Hence if $\beta_n{}^k/\alpha_n{}^k \geqq \frac{1}{2}$, then neither endpoint of $\gamma_n{}^k$ is a point of R_{n-2} and so, one is a point P_{n-1} and the other a point P_n.

Now suppose that $\liminf_{n\to\infty} [\beta_n(y)/\alpha_n(y)]$ is 1 on a set of positive measure. Then, since $\beta_n(y)/\alpha_n(y) \leqq 1$, it follows that $\beta_n(y)/\alpha_n(y) \to 1$ on a set of positive measure, and so there exists a y-set T of positive measure and a sufficiently large integer N such that $\beta_n(y)/\alpha_n(y) > \frac{1}{2}$, if $n \geqq N$ and y on T. Consider the points of T which lie in a particular $\gamma_N{}^k$. Then this interval $\gamma_N{}^k$ is divided into four intervals by a point P_{N+1} and two points P_{N+2}. The above considerations show that the points of T which are in $\gamma_N{}^k$ must lie in the two intervals determined by the point P_{N+1} and a point P_{N+2}. Each of these intervals is divided into two intervals by two points P_{N+3} and again the points of T which are in $\gamma_N{}^k$ must lie in the two intervals whose endpoints are one of these P_{N+3} and the corresponding P_{N+2}. Continuing in this manner, it is seen that there can be at most two points of T in any $\gamma_N{}^k$. This contradicts the fact that T is a set of positive measure.

3. The symmetric case. The set R will be called symmetric if the ratios $\alpha_n{}^k/\beta_n{}^k$ are independent of k, i. e., if $|\epsilon_n(y)| = \epsilon_n$ is independent of y. If the larger interval $\alpha_n{}^k$ is always nearer to the point $y = 0$ than the smaller

interval β_n^k, the set R will be called lower symmetric. In the symmetric case, formula (5) reduces to

$$\begin{aligned} &\epsilon_n(y) = +\epsilon_n \text{ if } y \text{ is in } \alpha_n(y) \\ &\epsilon_n(y) = -\epsilon_n \text{ if } y \text{ is in } \beta_n(y) \end{aligned} \qquad (n = 0, 1, \cdots), \tag{12}$$

where $0 \leqq \epsilon_n < 1$.

The case where ϵ_n is also independent of n has been treated by Faber [3], who has shown that in this case dy/dx can take only the two values 0 and $+\infty$. These simple functions seem to be the earliest examples of nowhere constant purely singular functions.

Next there will be proved

THEOREM III. *If R is symmetric, then $x = f^{-1}(y)$ is purely singular or absolutely continuous according as $\sum_{n=0}^{\infty} \epsilon_n^2$ is divergent or convergent.*

Since the inverse of a purely singular function is purely singular, Theorem III implies that $y = f(x)$ is purely singular if and only if the series $\sum_{n=0}^{\infty} \epsilon_n^2$ is divergent. On the other hand, the inverse of an absolutely continuous function is not necessarily absolutely continuous, i. e., it may be mixed. However, if R is symmetric, then $y = f(x)$ is absolutely continuous whenever $x = f^{-1}(y)$ is absolutely continuous. Thus the statement of Theorem III remains true if the function $x = f^{-1}(y)$ is replaced by the function $y = f(x)$. The proof of this fact depends on the theory of infinite convolutions and will be deferred to Section 6.

In order to prove Theorem III several lemmas will be needed.

LEMMA II. *If $\{\epsilon_n\}$ is an arbitrary fixed sequence for which $0 \leqq \epsilon_n < 1$, then the infinite product*

$$\prod_{n=0}^{\infty} (1 \pm \epsilon_n) \tag{13}$$

is convergent almost everywhere or divergent to zero almost everywhere according as $\sum_{n=0}^{\infty} \epsilon_n^2$ is convergent or divergent.

The "almost everywhere" of the lemma is meant in the usual sense (cf., e. g., Steinhaus [14]). A $(1, 1)$-correspondence is defined between the set of sequences $\{\pm, \pm, \cdots\}$ and the points of the interval $[0, 1]$, with the exception of an enumerable set on the latter interval. The sign $+$ is made to correspond to 0, the sign $-$ to 1. Then any sequence of $+$ and $-$ signs corresponds to the dyadic development of a point of the interval $[0, 1]$. Conversely, any point of the interval $[0, 1]$ (up to an enumerable set) has a unique

dyadic development which corresponds to a sequence of + and — signs. A set of sequences $\{\pm, \pm, \cdots\}$ is said to be measurable if the corresponding set of points in the interval $[0, 1]$ is measurable in the sense of Lebesgue; also the measure of a measurable set of sequences $\{\pm, \pm, \cdots\}$ is defined as the measure of its image set. Then the "almost everywhere," in the statement of Lemma II, means "for all sequences $\{\pm, \pm, \cdots\}$ with the possible exception of a set of sequences which has measure zero." [4]

Proof of Lemma II. The fact that the convergence of $\sum_{n=0}^{\infty} \epsilon_n^2$ implies the convergence of the product (13) almost everywhere is a well-known result of Rademacher [12]. However the entire lemma may be obtained with the help of a theorem of Khintchine and Kolmogoroff [6]. Clearly, the product (13) converges if and only if the corresponding sum $\sum_{n=0}^{\infty} \log(1 \pm \epsilon_n)$ converges. Now,

---

[4] It may be mentioned that the fact that the product (13), for a fixed sequence $\{\epsilon_n\}$, is either convergent almost everywhere or divergent almost everywhere is a consequence of the famous "0 or 1" principle (Cf. Kolmogoroff [8]). An elementary proof of this principle for the case under consideration will be given here. This proof may be extended to a more general class of problems.

Let $\{\delta_1, \delta_2, \cdots, \delta_n, \delta_{n+1}, \cdots\}$, where δ_n is a + sign or a — sign, be a particular sequence for which (13) converges. Then (13) also converges for any of the sequences $\{\pm, \pm, \cdots, \pm, \delta_{n+1}, \delta_{n+2}, \cdots\}$. Now let S be the set of points of the interval $[0, 1]$ which is the image of the set of all sequences $\{\pm, \pm, \cdots\}$ for which (13) converges. The set S is clearly measurable. Let $\psi(x)$, $0 \leqq x < 1$, denote the characteristic function of S, i. e., $\psi(x)$ is 1 or 0 according as x is or is not a point of S. Let the domain of definition of $\psi(x)$ be extended to the entire x-axis by the formula $\psi(x+1) = \psi(x)$. The above property of the image of S means that $\psi(x + (\frac{1}{2})^n) = \psi(x)$ for all x and for all integers n. Thus, $\psi(x)$ is a periodic, measurable function with arbitrarily small periods. Hence, $\psi(x)$ is constant almost everywhere. Thus, S is either a zero set or a set of full measure.

Added September 28, 1937. Using the method of Kac [*Commentarii Mathematici Helvetici*, vol. 9 (1936-37), pp. 170-171], it is possible to give a simple proof of the Burstin theorem [*Monatshefte für Mathematik und Physik*, vol. 26 (1915), p. 234] employed above. The theorem states that if $g(x)$ is a measurable function having arbitrarily small periods, then $g(x)$ is a constant almost everywhere. Let $\{\tau_n\}$ be a sequence of periods of $g(x)$ such that $\tau_n \to 0$, $n \to \infty$. The expression

$$M = \tau_n^{-1} \int_x^{x+\tau_n} e^{ig(t)} dt$$

is independent of both x and τ_n, since it is equal to

$$M(e^{ig(t)}) = \lim_{T \to \infty} T^{-1} \int_0^T e^{ig(t)} dt.$$

Now,

$$M = \tau_n^{-1} \int_x^{x+\tau_n} e^{ig(t)} dt \ (\to) \ e^{ig(x)}, \ n \to \infty;$$

so that, $g(x) \ (\equiv)$ const. $(\mathrm{mod}\ 2\pi)$. Application of this argument to $cg(x)$, where c is arbitrary, yields $g(x) \ (\equiv)$ const. $(\mathrm{mod}\ 2\pi/c)$. Hence $g(x) \ (=)$ const.

according to Khintchine and Kolmogoro$_3$, the sum $\sum_{n=0}^{\infty} \log(1 \pm \epsilon_n)$ converges almost everywhere if the mathematical expectancies

$$a_n = \tfrac{1}{2} \log(1 - \epsilon_n^2)$$

and the mean square deviations

$$b_n = \tfrac{1}{4} \left[\log \frac{1+\epsilon_n}{1-\epsilon_n} \right]^2$$

are such that $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ are convergent series; cf. also Jessen and Wintner [4], Theorems V and XXXII. These conditions are clearly satisfied if and only if the series $\sum_{n=0}^{\infty} \epsilon_n^2$ is convergent. Now it may be supposed that $|\log(1 \pm \epsilon_n)| < 1$, at least if $n > N$, since otherwise the sum $\sum_{n=N}^{\infty} \log(1 \pm \epsilon_n)$, and hence the product (13), is divergent everywhere. Under this assumption, a result of Kolmogoroff [7] implies that $\sum_{n=N}^{\infty} \log(1 \pm \epsilon_n)$ is divergent almost everywhere unless $\sum_{n=0}^{\infty} a_n$, $\sum_{n=0}^{\infty} b_n$ are convergent series, i. e., unless $\sum_{n=0}^{\infty} \epsilon_n^2$ is convergent. That the product (13) is divergent to zero almost everywhere if $\sum_{n=0}^{\infty} \epsilon_n^2 = +\infty$, may be seen from (15) below. This completes the proof of Lemma II.

LEMMA III. *If R is lower symmetric, then $x = f^{-1}(y)$ is purely singular or absolutely continuous according as $\sum_{n=0}^{\infty} \epsilon_n^2$ is divergent or convergent.*

Proof. Suppose y is a point of the interval $[0, 1]$ but not a point of the set R. Then y determines a unique product

$$\prod_{n=0}^{\infty} [1 + \epsilon_n(y)] = \prod_{n=0}^{\infty} [1 + \epsilon_n(f(x))] = \prod_{n=0}^{\infty} (1 \pm \epsilon_n). \tag{14}$$

Thus y determines uniquely a particular sequence $\{\pm, \pm, \cdots\}$, by (12). On the other hand, this sequence $\{\pm, \pm, \cdots\}$ corresponds to the dyadic development of a point x of the interval $[0, 1]$. The fact that R is lower symmetric clearly implies that this x satisfies $y = f(x)$. Thus the convergence or divergence of (13) "almost everywhere" in the sense of Lemma II means the convergence or divergence of $\prod_{n=0}^{\infty} [1 + \epsilon_n(f(x))]$ "almost everywhere" in

the usual sense, i. e., for all x in $[0, 1]$ with the possible exception of a zero set. Now, by (11) and (14),

$$\prod_{n=0}^{\infty} [1 + \epsilon_n(f(x))] (=) f'(x). \tag{15}$$

In virtue of (15), the proof of Lemma III is completed by an application of Lemma I and Lemma II.

Proof of Theorem III. Let x be a point of the interval $[0, 1]$ which does not possess a finite dyadic development. This x determines a unique y by the condition $y = f(x)$. Also, $y = f(x)$ determines a unique sequence $\{\pm, \pm, \cdots\}$, by (12). This sequence corresponds to the dyadic development of a point $\bar{x}$ of the interval $[0, 1]$. Thus, there is determined, up to an enumerable set, a measure preserving $(1, 1)$-transformation of the interval $[0, 1]$ into itself by the formula $x \to \bar{x}$. Now the symmetric set R determines a sequence of numbers $\{\epsilon_n\}$. Let $\bar{R}$ be the lower symmetric set which is uniquely determined by this sequence $\{\epsilon_n\}$ and let $y = \bar{f}(x)$, $0 \leqq x \leqq 1$, denote the corresponding function. Then,

$$\prod_{n=0}^{\infty} [1 + \epsilon_n(f(x))] (=) \prod_{n=0}^{\infty} [1 + \epsilon_n(\bar{f}(\bar{x}))]. \tag{16}$$

On the other hand, by (11),

$$f'(x) (=) \prod_{n=0}^{\infty} [1 + \epsilon_n(f(x))] \text{ and } \bar{f}'(x) (=) \prod_{n=0}^{\infty} [1 + \epsilon_n(\bar{f}(\bar{x}))]. \tag{17}$$

Thus, by (16) and (17),

$$f'(x) (=) \bar{f}'(\bar{x}).$$

So, by Lemma I and the fact that the correspondence $x \to \bar{x}$ is measure preserving, it is seen that the function $x = f^{-1}(y)$ is purely singular or absolutely continuous according as the function $\bar{x} = \bar{f}^{-1}(y)$ is purely singular or absolutely continuous. Theorem III now follows from Lemma III.

4. Nowhere monotone, absolutely continuous functions. Theorem III makes possible the construction, in a simple way, of functions which are absolutely continuous but are not monotone in any interval. Let R be a symmetric set associated with a sequence $\{\epsilon_n\}$, $0 \leqq \epsilon_n < 1$, for which $\sum_{n=0}^{\infty} \epsilon_n^2 < +\infty$ but $\sum_{n=0}^{\infty} \epsilon_n = +\infty$. By Theorem III, the function $x = f^{-1}(y)$ is absolutely continuous, where $y = f(x)$ is the function defined by R. Let $\gamma_m{}^k$ be an arbitrary interval determined by R_m. This $\gamma_m{}^k$ determines a set $(\delta_0, \delta_1, \cdots, \delta_m)$

of $+$ or $-$ signs, which are the first $(m+1)$ $+$ or $-$ signs determined by any y in $\gamma_m{}^k$. Now, since $\sum_{n=0}^{\infty} \epsilon_n = +\infty$, it is clear that there exist at least two sequences $\{\delta_0, \delta_1, \cdots, \delta_m, \pm, \pm, \cdots\}$ such that the product (13) diverges to $+\infty$ if the $\pm$ signs in the product are chosen to be those of one of these sequences and diverges to 0 if the $\pm$ signs are those of the other sequence. On the other hand, the reciprocal value of $\prod_{n=0}^{\infty} [1+\epsilon_n(y)]$ is a limit of difference quotients for the function $x=f^{-1}(y)$ at the point y. Thus in the interval $\gamma_m{}^k$ there exist at least two points such that the function $x=f^{-1}(y)$ has a derivative number equal to $+\infty$ at one of these points and at the other a derivative number equal to 0. Obviously, the function $x=f^{-1}(y)-y$ is absolutely continuous but not monotone in any y-interval.

In virtue of the remark following the statement of Theorem III, it may be seen that the function $y=f(x)-x$ is also absolutely continuous and nowhere monotone.

5. The modulus of continuity of purely singular functions. Let $y=f(x)$ be a function defined on the interval $[0,1]$; let

$$\omega(\delta) = \max |f(x_1) - f(x_2)|$$

for all x_1, x_2 belonging to $[0,1]$ and such that $|x_1 - x_2| \leqq \delta$. Then the function $\omega(\delta)$ is called the modulus of continuity of $f(x)$. The fact that $y=f(x)$ satisfies the Lipschitz condition of order α may be expressed by the formula $\omega(\delta) = O(\delta^\alpha)$. If $f(x)$ satisfies the Lipschitz condition of order one, $\omega(\delta) = O(\delta)$, then $f(x)$ is necessarily absolutely continuous. The object of the present section is to show that in a certain sense this statement cannot be improved.

For every monotone decreasing function $\phi(\delta)$ which approaches $+\infty$ as $\delta \to +0$, there exists a continuous, strictly increasing, purely singular function, $y=f(x)$, for which $\omega(\delta) = O(\delta\phi(\delta))$.

Proof. The function $y=f(x)$ will be chosen to be of the type considered in Section 3, i. e., it will be defined by the construction of a symmetric set R. In fact, every ϵ_n will be chosen to be either 0 or a fixed constant c, $0<c<1$. If ϵ_n is chosen in this manner, the set R will clearly be dense. Furthermore, by Theorem III, $f(x)$ will be purely singular if $\epsilon_n = c$ for infinitely many values of n.

Now it is clear that

$$\omega(2^{-n}) \leqq 2 \max_k \gamma_n^k = 2 \cdot 2^{-n} \prod_{j=0}^{n-1} (1 + \epsilon_j).$$

Thus we obtain

$$\omega(2^{-n})/2^{-n}\phi(2^{-n}) \leqq 2 \prod_{j=0}^{n-1} (1 + \epsilon_j)/\phi(2^{-n}). \tag{18}$$

Now, since $\phi(2^{-n})$ increases indefinitely with n, there exists a monotone sequence of positive integers $\{k_n\}$ such that

$$\phi(2^{-k_n}) > (1 + c)^n, \qquad (n = 0, 1, \cdots). \tag{19}$$

Now put $\epsilon_j = c$ or 0 according as j does or does not belong to the sequence $\{k_n\}$. Then

$$\prod_{j=0}^{n-1} (1 + \epsilon_j) = (1 + c)^i, \text{ where } k_i \leqq n - 1 < k_{i+1}. \tag{20}$$

Thus it follows from (18), (19), (20) and the monotony of $\phi(\delta)$ that

$$\omega(2^{-n})/2^{-n}\phi(2^{-n}) \leqq 2(1 + c)^i/\phi(2^{-k_i}) \leqq 2. \tag{21}$$

Now let $2^{-(n-1)} < \delta \leqq 2^{-n}$, then $\omega(\delta) \leqq \omega(2^{-n})$. Also

$$\omega(\delta)/\delta \leqq [\omega(2^{-n})/2^{-n}][2^{-n}/\delta] \leqq 2[\omega(2^{-n})/2^{-n}]. \tag{22}$$

On the other hand, $\phi(\delta) > \phi(2^{-n})$, so that, by (21) and (22),

$$\omega(\delta)/\delta\phi(\delta) \leqq 2\omega(2^{-n})/2^{-n}\phi(2^{-n}) \leqq 4.$$

This completes the proof of the italicized statement above.

6. Certain Poisson convolutions. Let $\sigma_n = \sigma_n(x)$ denote the distribution function

$$\begin{aligned} &\sigma_n(x) = 0, && -\infty < x \leqq 0; \\ &\sigma_n(x) = \tfrac{1}{2}(1 + \epsilon_n), && 0 < x \leqq (\tfrac{1}{2})^{n+1}; \\ &\sigma_n(x) = 1, && (\tfrac{1}{2})^{n+1} < x < +\infty; \end{aligned} \tag{23}$$

$0 \leqq \epsilon_n < 1$; $n = 0, 1, \cdots$. Thus $\sigma_n(x)$ is the step function with the jump $\frac{1}{2}(1 + \epsilon_n)$ at $x = 0$ and the jump $\frac{1}{2}(1 - \epsilon_n)$ at $x = (\frac{1}{2})^{n+1}$. Clearly, the convolution

$$\tau_1(x) = \sigma_0 * \sigma_1 = \int_{-\infty}^{+\infty} \sigma_0(x - \xi)\, d\sigma_1(\xi)$$

is the step function with the jumps

$$\begin{aligned} &(\tfrac{1}{2})^2(1 + \epsilon_0)(1 + \epsilon_1) \text{ at } x = 0, \\ &(\tfrac{1}{2})^2(1 + \epsilon_0)(1 - \epsilon_1) \text{ at } x = (\tfrac{1}{2})^2, \\ &(\tfrac{1}{2})^2(1 - \epsilon_0)(1 + \epsilon_1) \text{ at } x = 2(\tfrac{1}{2})^2, \text{ and} \\ &(\tfrac{1}{2})^2(1 - \epsilon_0)(1 - \epsilon_1) \text{ at } x = 3(\tfrac{1}{2})^2. \end{aligned}$$

Similarly, the convolution

$$\tau_n(x) = \sigma_0 * \sigma_1 * \cdots * \sigma_n = \tau_{n-1} * \sigma_n \tag{24}$$

has jumps $(\frac{1}{2})^{n+1} \prod_{j=0}^{n} (1 \pm \epsilon_j)$ at the points $k/2^{n+1}$, $k = 0, 1, \cdots, 2^{n+1} - 1$. The arrangement of $+$ and $-$ signs in the product which represents the jump of τ_n at the point $k/2^{n+1}$ is the same as the arrangement of zeros and ones in the finite dyadic development of the point $k/2^{n+1}$. Let $\gamma^k{}_{n+1}$ denote the jump

$$\gamma^k{}_{n+1} = (\tfrac{1}{2})^{n+1} \prod_{j=0}^{n} (1 \pm \epsilon_j), \qquad (k = 0, 1, \cdots, 2^{n+1} - 1) \tag{25}$$

of τ_n at the point $x = k/2^{n+1}$. Let

$$y^k{}_{n+1} = \sum_{j<k} \gamma^j{}_{n+1}, \qquad (k = 0, 1, \cdots, 2^{n+1}), \tag{26}$$

so that

$$\tau_n(k/2^{n+1}) = y^k{}_{n+1}. \tag{27}$$

Now, by (25) and the remark concerning the arrangement of $+$ signs and $-$ signs in (25), it is clear that

$$\gamma^{2k}_{n+2} = \tfrac{1}{2}\gamma^k{}_{n+1}(1 + \epsilon_{n+2}) \quad \text{and} \quad \gamma^{2k+1}_{n+2} = \tfrac{1}{2}\gamma^k{}_{n+1}(1 - \epsilon_{n+2}),$$

so that

$$\gamma^{2k}_{n+2} + \gamma^{2k+1}_{n+2} = \gamma^k{}_{n+1}.$$

Thus, by (26),

$$y^k{}_{n+1} = y^{2k}_{n+2}. \tag{28}$$

From (27) and (28) it is seen that

$$\tau_m(k/2^{n+1}) = y^k{}_{n+1}, \qquad (m \geqq n). \tag{29}$$

Now let $\tau(x)$ denote the infinite convolution

$$\tau(x) = \sigma_0 * \sigma_1 * \cdots = \lim_{n\to\infty} \tau_n(x). \tag{30}$$

The conditions for the convergence of this infinite convolution are satisfied in the case under consideration for any choice of $\{\epsilon_n\}$, $0 \leqq \epsilon_n < 1$; cf. Jessen and Wintner [4], Theorem V. It will be supposed that

$$(\tfrac{1}{2})^{n+1} \prod_{j=0}^{n} (1 + \epsilon_j) \to 0, \qquad (n \to \infty). \tag{31}$$

This assures the continuity of $\tau(x)$. Now, comparing (29) with (30),

$$\tau(k/2^{n+1}) = y^k{}_{n+1}. \tag{32}$$

Thus, by (1), the set of points $y^k{}_{n+1}$ is the section R_{n+1} of the set R associated with $\tau(x)$. Furthermore, by (26), the $\gamma^k{}_{n+1}$ defined by (25) are exactly the intervals $\gamma^k{}_{n+1}$ determined by R_{n+1} (cf. (3)). Comparing (10), (12), (25) and the definition of a lower symmetric set R, it is seen that the set R associated with the function (30) is lower symmetric. Conversely, any continuous, strictly increasing function associated with a lower symmetric set R is the infinite convolution of Poisson distribution functions (23).

It has been shown by Jessen and Wintner ([4], Theorem XXXV) that a continuous infinite convolution of step functions is either purely singular or absolutely continuous. Thus a function $y = \tau(x)$ for which the corresponding set R is lower symmetric is either purely singular or absolutely continuous. On the other hand, it has been shown in Section 3, Lemma III, that $y = \tau(x)$ is purely singular if and only if $\sum_{n=0}^{\infty} \epsilon_n{}^2 = +\infty$. It follows that $y = \tau(x)$ is absolutely continuous if and only if $\sum_{n=0}^{\infty} \epsilon_n{}^2 < +\infty$.

By the use of a measure preserving transformation on the y-axis similar to that introduced in the proof of Theorem III, this statement can be extended to apply to the general symmetric case.

THEOREM IV. *If R is symmetric, then $y = f(x)$, together with $x = f^{-1}(y)$, is purely singular or absolutely continuous according as $\sum_{n=0}^{\infty} \epsilon_n{}^2$ is divergent or convergent.*

It should be noted that, in the definition (23) of $\sigma_n(x)$, it was not essential to restrict ϵ_n to be positive. In fact, let $\rho_n(x)$ be any distribution function of the form

$$
\begin{aligned}
&\rho_n(x) = 0, && -\infty < x \leqq 0; \\
&\rho_n(x) = q_n, && 0 < x \leqq (\tfrac{1}{2})^{n+1}; \\
&\rho_n(x) = 1, && (\tfrac{1}{2})^{n+1} < x < +\infty;
\end{aligned} \tag{33}
$$

$0 < q_n < 1$; $n = 0, 1, \cdots$. Then, if the infinite convolution $\rho_0 * \rho_1 * \cdots$ is continuous, it is of the type considered in Theorem IV. Thus

THEOREM V. *The infinite Poisson convolution $\rho_0 * \rho_1 * \cdots$ of distribution functions (33) is continuous if and only if $\prod_{j=0}^{n} \max(q_j, 1 - q_j) \to 0$, $n \to +\infty$. If continuous, it is purely singular or absolutely continuous, along with its inverse, according as $\sum_{j=0}^{\infty} (q_j - \frac{1}{2})^2$ is divergent or convergent.*

THE JOHNS HOPKINS UNIVERSITY,
THE UNIVERSITY OF WISCONSIN.

BIBLIOGRAPHY.

[1] T. Brodén, "Beiträge zur Theorie der stetigen Functionen einer reellen Veränderlichen," *Journal für die reine und angewandte Mathematik*, vol. 118 (1897), pp. 1-60.

[2] A. Denjoy, "Sur une fonction de Minkowski," *Comptes Rendus* (Paris), vol. 198 (1934), pp. 44-47.

———, "Sur quelques points de la théorie des fonctions," *Comptes Rendus* (Paris), vol. 194 (1932), pp. 44-46.

[3] G. Faber, "Über stetige Funktionen II," *Mathematische Annalen*, vol. 69 (1910), pp. 372-443; cf. pp. 395-400.

[4] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 44-88.

[5] E. R. van Kampen and A. Wintner, "On a singular monotone function," *Journal of the London Mathematical Society* (in press).

[6] A. Khintchine and A. Kolmogoroff, "Über Konvergenz von Reihen, deren Glieder durch den Zufall bestimmt werden," *Recueil Mathématique de Moscou*, vol. 32 (1925), pp. 668-677.

For another treatment and for further results, see Kolmogoroff [7].

[7] A. Kolmogoroff, "Über die Summen durch den Zufall bestimmter unabhängiger Grössen," *Mathematische Annalen*, vol. 99 (1928), pp. 309-319 and vol. 102 (1930), pp. 484-488.

[8] A. Kolmogoroff, "Grundbegriffe der Wahrscheinlichkeitsrechnung," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 2, Heft 3; in particular, see the appendix.

[9] H. Lebesgue, *Leçons sur l'intégration*, 2nd edition. In particular, see chap. IX, section 1.

[10] H. Minkowski, "Zur Geometrie der Zahlen," *Verhandlungen des III. Internationalen Mathematiker-Kongresses in Heidelberg*, 1904, pp. 164-173.

[11] H. Rademacher, "Zu dem Borelschen Satz über die asymptotische Verteilung der Ziffern in Dezimalbrüchen," *Mathematische Zeitschrift*, vol. 2 (1918), pp. 306-311; cf. also vol. 3 (1919), p. 317.

[12] H. Rademacher, "Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen," *Mathematische Annalen*, vol. 87 (1922), pp. 112-138.

[13] W. Sierpiński, "Un exemple élémentaire d'une fonction croissante qui a presque partout une dérivée nulle," *Giornale de Matematiche di Battaglini*, vol. 54 (1916), pp. 314-334.

[14] H. Steinhaus, "Les probabilités dénombrables et leur rapport à la théorie de la mesure," *Fundamenta Mathematicae*, vol. 4 (1923), pp. 286-310.

[15] A. Wintner, "On convergent Poisson convolutions," *American Journal of Mathematics*, vol. 57 (1935), pp. 827-838.

PLANAR GRAPHS WHOSE HOMEOMORPHISMS CAN ALL BE EXTENDED FOR ANY MAPPING ON THE SPHERE.*

By V. W. Adkisson and Saunders MacLane.

Introduction. Certain conditions have been found under which a given homeomorphism of a given point set on a sphere can be extended to the sphere.[1]

It is the purpose of this paper to characterize the cyclicly connected planar graphs[2] G such that in every map of G on the sphere every homeomorphism of G into itself can be extended to the sphere.[3] The characterization will use only internal properties of the graph G, defined in G without reference to a map of G on any other space.

It is known that the extendability of a homeomorphism of a map of G depends on the behavior of the complementary domain boundaries of the map under the homeomorphism. Therefore the procedure for our problem is a study to determine which simple closed curves of a given graph can be boundaries of complementary domains of a map of a graph.

Definitions. The circuit[4] J of a cyclicly connected graph G is called a *bounding circuit* of G provided that for any two distinct maximal connected components, N_1 and N_2, of $G - J$ the sets $\bar{N}_1 \cdot J$ and $\bar{N}_2 \cdot J$ lie respectively on two distinct arcs AXB and AYB of J.[5]

---

* Presented to the Society, December 30, 1936. Received by the Editors February 6, 1937; revised April 8, 1937.

[1] H. M. Gehman, "On extending a homeomorphism between two subsets of spheres," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 79-81; V. W. Adkisson, "Cyclicly connected continuous curves whose complementary domain boundaries are homeomorphic," *Comptes Rendus des séances de la Société des Sciences et des Lettres de Varsovie*, vol. 23 (1930), Classe III, pp. 164-193. (Referred to in this paper as Thesis); V. W. Adkisson, "On extending a continuous (1 — 1) correspondence of continuous curves on a sphere," *ibid.*, vol. 27 (1934).

[2] For a discussion of cyclicly connected curves in general see G. T. Whyburn, "Cyclicly connected continuous curves." *Proceedings of the National Academy of Sciences*, vol. 13 (1927), pp. 31-38. A cyclicly connected graph is also non-separable. See H. Whitney, "Non-separable and planar graphs," *Transactions of the American Mathematical Society*, vol. 34 (1932), p. 339.

[3] This paper is the outgrowth of a problem originally suggested by Professor J. R. Kline.

[4] A circuit is a simple closed curve.

[5] "Bounding circuit" is equivalent to "*boundary curve*" as defined by S. Claytor,

Note that each set $\bar{N}_i \cdot J$ consists of only a finite number of vertices of G. These vertices may be called "feet" of N_i.

A *split-circuit* of G is a bounding circuit J such that $G - J$ contains at least two components. The chief result of this paper may be stated in the following theorem.

THEOREM C. *Every homeomorphism of the cyclicly connected graph G into itself is extendable in every map of G on the sphere if, and only if, for each split circuit J and every homeomorphism σ of G such that $\sigma(G) = G$ and $\sigma(J) \neq J$, the sets $(G - J)$ and $J \cdot \sigma(J)$ each consists of two maximal connected components.*

The final theorem (Theorem D) characterizes more precisely the combinatorial structure of the graphs G satisfying the conditions of Theorem C. Specifically, such a graph either is a graph with just one map on the sphere or else is a graph with at most four essentially distinct homeomorphisms.

LEMMA 1. *Let the vertices p and q of a graph G lie on a circuit J of G, and in a map of G on a sphere S let R be one of the regions of S bounded by J. Then there is a complementary domain[6] boundary of G which contains both p and q and which lies in $\bar{R}$ if, and only if, there is no open arc of G in R which has ends on J and whose ends separate p and q on J.*[7]

THEOREM A. *If J is a circuit of G, then G can be mapped on a sphere so that J is a complementary domain boundary of G if, and only if, J is a bounding circuit.*[8]

This theorem can be proved by the following procedure. Take any map of G in which a certain component of $G - J$, say N, lies "outside" J.[9] One

---

"Topological immersion of peanian continua in a spherical surface," *Annals of Mathematics*, vol. 35 (1934), p. 809. A simple closed curve J of G is called a boundary curve of G provided that there do not exist in $G - J$ distinct components N_1 and N_2 such that (1) a point-pair of $\bar{N}_1 \cdot J$ separates a point pair of $\bar{N}_2 \cdot J$ on J, or (2) $\bar{N}_1 \cdot J = \bar{N}_2 \cdot J =$ three distinct points. Both definitions will be found useful in later proofs.

[6] A "complementary domain" will refer always to a domain complementary to some given map of G on a sphere. "Complementary domain boundary" will be abbreviated as "c. d. b."

[7] This lemma follows from a lemma by Kuratowski, "Sur les courbes gauches," *Fundamenta Mathematicae*, vol. 15 (1930), p. 274, Lemma III'; it can also be proved directly for graphs by more elementary combinatorial arguments.

[8] This is Proposition K of Claytor, *loc. cit.*, p. 828.

[9] It is sometimes convenient to refer to the two regions of a sphere bounded by a circuit J as "outside" and "inside" of J.

is able to show then, from the definition of a split-circuit, that the feet of N all lie on one c. d. b. of G "inside" and on J. Hence the given map can be altered by mapping N inside J instead of outside J. Proceeding in this manner we obtain a new map of G with all of G inside and on J.

COROLLARY 1. *If there is only one component of $G - J$ the circuit J is a c. d. b. of any map of G on a sphere.*

COROLLARY 2. *If J is a split circuit, G can be mapped so that J is not a c. d. b. of G.*

COROLLARY 3. *The c. d. b. J of a map of G will have a disconnected intersection with some other c. d. b. of the map of G if, and only if, J is a split circuit.*

Proof. The condition is sufficient. If J is a split circuit there are at least two components, N_1 and N_2, of $G - J$, while J is composed of two arcs AXB and AYB, such that $\bar{N}_1 \cdot J$ lies in AXB and $\bar{N}_2 \cdot J$ in AYB. Since G is cyclic, $\bar{N}_1 \cdot J$ is not a point, and A and B can be so chosen as to both belong to $\bar{N}_1 \cdot J$. Let R be the complementary domain of G bounded by J and set $R_1 = S - \bar{R}$. Now A and B both lie on a c. d. b. K of G with $K \subset \bar{R}_1$. For if not, by Lemma 1 there would exist a component N of $G - J$ such that two points of $\bar{N} \cdot J$ separate the points A and B on J and J would not be a split circuit. Since there is a component N_2 of $G - J$ such that $\bar{N}_2 \cdot J \subset AYB$, the arc AYB cannot be a subset of K, except possibly when $\bar{N}_2 \cdot J = A + B$. In this case a different choice of K readily gives the result. In like manner the arc AXB is not a subset of K. Therefore, since A and B are common to both K and J, $K \cdot J$ is disconnected.

The condition is necessary. Suppose G is mapped so that the c. d. b. J of G and the c. d. b. J_1 of G have a disconnected intersection F. Let F_1 and F_2 be two points in separate maximal connected subsets of this intersection. Then the removal of F_1 and F_2 disconnects G. For let the open arc $\langle F_1XF_2\rangle$ be any arc lying in the domain complementary to G bounded by J and $\langle F_1YF_2\rangle$ an arc lying in the domain complementary to G bounded by J_1. The simple closed curve $F_1XF_2YF_1$ has only the points F_1 and F_2 in common with G and disconnects G. Then F_1 and F_2 separate G into two components and neither component consists of a single arc, for otherwise F_1 and F_2 would belong to the same connected subset of F, contrary to assumption. Therefore there must exist two or more components of $G - J$, and since J is a c. d. b. of G, Theorem A shows J to be a bounding circuit. Then J is a split circuit.

THEOREM B. *In any mapping of G on a sphere S there are two complementary domain boundaries of G with a disconnected intersection if, and only if, G contains a split circuit.*

Proof. The condition is necessary from Corollary 3.

The condition is sufficient. Suppose G contains a split circuit J. Let N_1 and N_2 be two components of $G - J$ and divide J relative to N_1 and N_2 into two arcs AXB and AYB with A and B in $\bar{N}_1 \cdot J$, as in the proof of Corollary 3. Since J is a split circuit, every component M of $G - J$ satisfies either $\bar{M} \cdot J \subset AXB$ or $\bar{M} \cdot J \subset AYB$. Hence $G - (A + B)$ consists of two mutually separated sets; $H_1 = \langle AXB \rangle + N_1$ plus all components M, except N_1, with $\bar{M} \cdot J \subset AXB$, and $H_2 = \langle AYB \rangle + N_2$ plus all remaining components of $G - J$. Arcs of both sets H_1 and H_2 end on A, and so among the c. d. b.'s of G passing through A there will be at least one c. d. b. K containing arcs in both H_1 and H_2. Thus K must pass through A and B, the only points common to $\bar{H}_1$ and $\bar{H}_2$. Since A and B disconnect G into H_1 and H_2, neither one an arc, K also disconnects G, so that K is a split circuit. It is also a c. d. b. of G, so that by Corollary 3 it has a disconnected intersection with some other c. d. b., as asserted.

Proof of Theorem C. The conditions are necessary. First suppose $\sigma(J)$ does not contain any points of a component N_1 of $G - J$. Take a map of G in which J is a c. d. b. of G. As σ is extendable, $\sigma(J)$ must also be a c. d. b. of this map. We can map N_1 in the complementary domain of G bounded by J and leave $G - N_1$ fixed. But this gives a new map of G in which J is not a c. d. b. while $\sigma(J)$ remains a c. d. b. But then σ^{-1} cannot be extended to the sphere,[10] contrary to assumption. Hence $\sigma(J)$ contains points in every component of $G - J$.

Since there are at least two components N_1 and N_2 of $G - J$, $\sigma(J)$ to pass from N_1 to N_2 must contain points in J and also points not in J. Let $J \cdot \sigma(J)$ consist of k connected pieces, $g_1, g_2, \cdots, g_k$, each piece an arc or a point. Then $J - J \cdot \sigma(J)$ consists of k arcs $F_1, F_2, \cdots, F_k$, where the notation is chosen so that

$$J = F_1 + g_1 + F_2 + g_2 + \cdots + F_k + g_k \tag{1}$$

in the cyclic order shown. Because J and $\sigma(J)$ are both c. d. b.'s of G, it follows readily that the points of $g_1, \cdots, g_k$ are arranged in the same cyclic

[10] Thesis, Theorem 2. If M is a cyclicly connected continuous curve lying on a sphere S, and T a homeomorphism such that $T(M) = M$, a necessary and sufficient condition that T be extendable to S is that for every c. d. b. J of M, $T(J)$ be also a c. d. b.

order on $\sigma(J)$ as on J. Hence the arcs of $\sigma(J) - J \cdot \sigma(J)$ can be so denoted by $f_1, f_2 \cdots f_k$ that the cyclic order of arcs on $\sigma(J)$ is

$$\sigma(J) = f_1 + g_1 + f_2 + g_2 + \cdots + f_k + g_k. \tag{2}$$

The ends of f_i can then be so denoted by A_i and B_i that these points appear on J and on $\sigma(J)$ in the cyclic order $A_1B_1A_2B_2 \cdots A_kB_k$.

The circuits J and $\sigma(J)$ together divide the sphere into $k+2$ regions with the boundaries J, $\sigma(J)$, and $f_i + F_i$ $(i = 1, \cdots, k)$. As J and $\sigma(J)$ are c. d. b.'s of the map of G, the rest of G lies in the regions bounded by $f_i + F_i$. But each component of $G - J$ must contain some points of $\sigma(J)$ and hence some arc f_i. Therefore each closed region bounded by an $f_i + F_i$ contains exactly one component N_i of $G - J$, and there are k such components. Because σ carries J into $\sigma(J)$, $G - \sigma(J)$ must also have k components; namely, $N_i + F_i - (f_i + A_i + B_i)$.

Suppose now that there were three (or more) components of $G - J$. We can map N_1 in the complementary domain of G bounded by J, leaving the rest of G unaltered. This simply interchanges f_1 and F_1 in the boundaries J and $\sigma(J)$, so that there results a map of G with the c. d. b.,

$$K = f_1 + F_2 + \cdots + F_k + (J \cdot (\sigma J)), \quad L = F_1 + f_2 + \cdots + f_k + (J \cdot (\sigma J)).$$

These are the only c. d. b. containing $J \cdot \sigma(J)$, because every other c. d. b. is contained in one of $N_i + F_i$. Since σ must also be extendable in this map, it must carry K into K or L. But σ carries J into $\sigma(J)$, hence carries each F_i into some f_j, and carries K into a c. d. b. containing $\sigma(f_1)$, $k-1$ of the f_j, and $J \cdot (\sigma J)$. As this c. d. b. can be only L, and F_1 is in L, we have $\sigma(f_1) = F_1$. By the same argument, $\sigma(f_i) = F_i$ for every i.

But σ must leave fixed each end A_1 and B_1 of f_1, for otherwise $\sigma(A_1) = B_1$, $\sigma(B_1) = A_1$ which is impossible since $\sigma(A_i + B_i) = A_i + B_i$ $(i = 1, 2, \cdots, k)$ and $k \geqq 3$. This contradiction shows $\sigma A_1 = A_1$, $\sigma B_1 = B_1$.

Then also $\sigma f_1 = F_1$, $\sigma(F_1) = f_1$ and σ must carry $N_1 + F_1$ into itself. There is then another homeomorphism σ', equal to σ on N_1 and to the identity elsewhere. This σ' carries the c. d. b. J into $(J - F_1) + f_1$, which is not a c. d. b., so that σ' is not extendable in this map, contrary to hypothesis. The assumption of more than two components of $G - J$ is thus inconsistent. Therefore $G - J$ consists of two pieces N_1 and N_2, and we have already shown that $J \cdot \sigma(J)$ consists of two connected pieces, namely two arcs (or points) B_1A_2 and B_2A_1. Therefore the conditions of the theorem are necessary.

The conditions are sufficient. We need only show for each bounding

circuit J that in any map in which J is a c. d. b. of G, $\sigma(J)$ must be a c. d. b. of G. (Thesis, Theorem 2). Consider a map in which J is a c. d. b. of G. If J is not a split circuit then $\sigma(J)$ is not a split circuit and must be a c. d. b. of G. (Corollary 1, Theorem A). If $\sigma(J) = J$ the case is trivial. Hence, assume $\sigma(J) \neq J$. Denote the two components of $J \cdot \sigma(J)$ by α and β. Then J consists of four parts F_1, α, F_2, β in that cyclic order, where F_1 and F_2 are arcs, α and β may be arcs or points. The circuit $\sigma(J)$ consists of four arcs f_1, α, f_2, β in that cyclic order, where f_i and F_i have end points in common. Hence $J + \sigma(J)$ divides the sphere into four regions, r_1, r_2, r_3, r_4, with the respective boundaries: J, $\sigma(J)$, $f_1 + F_1$, and $f_2 + F_2$.

Now the arcs F_1 and F_2 of J belong to distinct components of $G - \sigma(J)$. For otherwise there would be an arc f of $G - \sigma(J)$ joining F_1 to F_2, and it readily follows that part of this arc f must lie in the region r_1, contrary to the fact that r_1 is a complementary domain of G.

Suppose now that $\sigma(J)$ is not a c. d. b. of G. Then there are arcs of G in r_2 bounded by $\sigma(J)$, so that at least one component of $G - \sigma(J)$ is contained in this region. But there are at least two other components of $G - \sigma(J)$ (those containing F_1 and F_2) not in r_2. Hence there are at least three components of $G - \sigma(J)$. But this is impossible since there are only two components of $G - J$ by hypothesis, and σ carries components of $G - J$ into components of $G - \sigma(J)$. Therefore, $\sigma(J)$ is a c. d. b. of G and the theorem is proved.

We shall now obtain other necessary and sufficient conditions that every homeomorphism of G be extendable in every map. These new conditions will give in some detail a more complete characterization of the graphs G.

LEMMA 2. *If every homeomorphism of G into itself is extendable in every map on the sphere, and if for some homeomorphism σ, some split circuit J and some component N_1 of $G - J$, $\sigma(J) = J$ and $\sigma(N_1) = N_1$, then σ has two fixed points on J.*

Proof. Because J is a split circuit we can choose an arc $F = AYB$ of J such that F contains only A and B from $\bar{N}_1 \cdot J$, while F contains $\bar{N}_2 \cdot J$ for some component N_2 of $G - J$. Map G with N_1 inside J and all other components of $G - J$ outside J. There is then inside J a domain complementary to G whose boundary K contains F. Furthermore $K = F + f$, where f is an open arc of N_1.

Suppose first that $\sigma(K) = K$. Then $\sigma(J) = J$ implies that $\sigma(f) = f$, $\sigma(F) = F$ and hence that $\sigma(A + B) = A + B$. If A and B remain fixed

we have the two desired fixed points. If A and B are interchanged by σ, there is a fixed point on [11] $F = AYB$ and also a fixed point on $J - F$.

Now assume $\sigma(K) \neq K$. The component N_2 is separated from $G - N_2$ by F and hence by $K = F + f$. Therefore K is a split circuit and N_2 a component of $G - K$. Theorem C applied to K shows that $\sigma(K)$ contains an arc of N_2. But this is impossible since $\sigma(K)$ consists of an arc of $\sigma(J) = J$ and an arc of $\sigma N_1 = N_1$, neither in N_2. Therefore $\sigma(K) = K$ and the lemma is proved.

Two homeomorphisms σ and τ are considered identical if $\sigma(G) = G$, $\tau(G) = G$ such that branch points that correspond under σ also correspond under τ, and edges that correspond under σ correspond under τ. (An edge of G means an arc of G joining two branch points, but containing no other branch points.)

THEOREM D. *Every homeomorphism of a cyclicly connected graph G into itself is extendable in every map of G on the sphere if, and only if, at least one of the following conditions holds:*

I. *Every split circuit of G is invariant under every homeomorphism σ. If there is a split circuit, then there is at most one $\sigma \neq 1$.*

II. *No split circuit is invariant under any $\sigma \neq 1$; there is only one homeomorphism $\sigma \neq 1$ and this σ is of order 2. For any split circuit J, $G - J$ consists of two connected pieces N_1 and N_2; $J \cdot \sigma(J)$ consists of just two connected pieces M_1 and M_2, and $\sigma(N_1) =$ a subset of N_1 plus an arc of J, $\sigma(N_2) =$ a subset of N_2 plus an arc of J, $\sigma(M_1) = M_2$, $\sigma(M_2) = M_1$.*

III. *There are in G only four split circuits, J_1, J_2, J_3, J_4 and only two topologically distinct maps of G on the sphere are possible. The complementary domain boundaries of each map include exactly two split circuits: J_1 and J_2 in one map, and J_3, J_4 in the other map. Any σ is of order 2 and either carries every split circuit into itself or else interchanges both J_1 and J_2, and J_3 and J_4. There are at most four distinct σ's, including the identity.*

Proof. Suppose every homeomorphism of G extendable.

Case 1. $\sigma(J) = J$ for every homeomorphism σ and every split circuit J. If some σ leaves some split circuit J pointwise fixed then σ is the identity.

[11] W. L. Ayres, "Concerning continuous curves and correspondences," *Annals of Mathematics*, vol. 28 (1927), p. 402.

For, map G so that J is a c. d. b. of G, and extend σ to be a periodic homeomorphism σ' of the sphere.[12] Since this σ' has a circuit of fixed points Eilenberg's results [13] concerning periodic transformations of a sphere show that the extension σ' is either the identity or is homeomorphic to a reflection in J. In the latter case σ' would take $G - J$, on one side of J, into something on the other side of J; but there is nothing on the other side of J since J is a c. d. b. of G. Hence σ must be the identity, as asserted.

Consider now any split circuit J and any component N_1 of $G - J$. As in the proof of Lemma 2 choose a split circuit $K = f + F$, where $f \subset N_1$ and $F \subset J$. For any σ, the hypothesis of Case 1 shows that $\sigma(K) = K$, $\sigma(J) = J$ and hence that $\sigma(F) = F$. Let A and B be the ends of the arc F. If A and B are each invariant under σ, then every edge and branch point of J is fixed under σ, and σ is the identity. The only alternative is for σ to interchange A and B.

If there were two homeomorphisms σ and τ, neither the identity, then both σ and τ must interchange A and B. Since $\sigma \cdot \tau^{-1}$ leaves A and B fixed, it therefore will be the identity, so that $\sigma = \tau$. Hence there is, as asserted in I, only one homeomorphism not the identity.

Case 2. For some split circuit J, and some σ, $\sigma(J) \neq J$. By Theorem C there are then only two components N_1 and N_2 of $G - J$. Introduce the subarcs F_1 and F_2 of J and the subarcs f_1 and f_2 of $\sigma(J)$ as in (1) and (2) in the proof of Theorem C, and consider the four split circuits,

$$(3) \qquad J,\ \sigma(J),\ F_1 + f_2 + J \cdot \sigma(J),\ f_1 + F_2 + J \cdot \sigma(J),$$

composed of J, f_1 and f_2. We shall show first that if G contains a split circuit K, other than one of (3), then K must lie entirely in one of the sets $H_i = N_i + F_i$ $(i = 1, 2)$.

Suppose G contains a split circuit K that is not one of (3) and has points in both H_1 and H_2. Let g_i be the arc of K that lies in H_i and A and B the ends of F_i. Then the circuits K and L $(= K - g_2 + F_2)$ each pass through A and B. Now one of the arcs, say g_1, must be different from F_1 and from f_1, and L is a split circuit; for any two components of $G - L$ with feet on the arc $L - F_2$ must have feet on distinct arcs (except for end points) of L since

[12] Since G is a graph (consisting of a finite number of vertices and edges) any transformation σ such that $\sigma(G) = G$ is necessarily periodic with a finite period k. Hence, if σ is extendable to the sphere, a transformation σ' of the sphere into itself can be so defined that $\sigma' = \sigma$ for points of G and that σ' is periodic of period k.

[13] S. Eilenberg, "Sur les transformations periodique de la surface de sphere," *Fundamenta Mathematicae* vol. 22 (1934), pp. 28-41.

they are on distinct arcs of the split circuit K, and there is only one component N_2 of $G - L$ with feet on F_2. If G is mapped so that J and $\sigma(J)$ are both c. d. b.'s of the map of G, the region R inside $f_1 + F_1$ (i. e. the region not containing F_2) is cut by the arc g_1. Since $\sigma(J) \neq J$ it follows easily that $\sigma(L) \neq L$, and from Theorem C, $G - L$ has only two components. But since the region R is cut by the arc g_1 of L there must be two components of $G - L$ in $\bar{R}$ and one component N_2 in H_2, making three in all. Therefore, no split circuit except those in (3) has arcs in both H_1 and H_2.

Suppose $\sigma(H_i) = H_i$. Then $\sigma(J) \neq J$, $\sigma(f_i) = F_i$ and $\sigma(F_i) = f_i$. Furthermore, if the end points of F_i remain invariant under σ it is possible to define a homeomorphism τ equal to σ on H_1 and to the identity on $G - H_1$. Then τ is not extendable (Theorem C) contrary to assumption. Hence the end points of F_i are interchanged by σ. Suppose now that H_1 contains a split circuit K such that $\sigma(K) = K$. Then K is in R_1, the "inside" of $f_1 + F_1$. Let R_2 be the region "outside" $f_1 + F_1$. Extend σ to be a periodic transformation σ' of the sphere. Since H_2 lies in R_2, $\sigma'(R_2) = R_2$. The periodic transformation σ' of R_2 must be homeomorphic to a rotation or to reflection [14] of $\bar{R}_2$. The latter is impossible as σ' has no fixed points on the boundary of R_2. The "rotation" σ' thus has a fixed point in R_2, and by Lemma 2 σ' has two fixed points in R_1. But $f_1 + F_1$ is fixed under σ' so that σ' must be homeomorphic to a rotation of the sphere, although a rotation of the sphere cannot have a fixed circuit and three fixed points not on this circuit (Eilenberg). Hence $\sigma(K) \neq K$ so that G contains no split circuit invariant under σ. Therefore, if there is no homeomorphism of G sending H_1 into H_2, σ is the only homeomorphism not the identity. This follows as in the last paragraph of Case 1. Then $\sigma^2 = 1$,[15] and we have condition II.

The only possibility remaining is that some σ has $\sigma(J) \neq J$ with $\sigma(H_1) = H_2$ and $\sigma(H_2) = H_1$. Suppose a split circuit K lies in H_1. Then $\sigma(K) \neq K$ and there are only two components of $G - K$. One of these components must contain H_2, and the other must lie in H_1. Now $\sigma(K)$ must contain an arc in each of these components, and hence must have an arc in H_1. But this is impossible since $\sigma(K)$ must lie in H_2. Therefore, G contains no split circuits other than (3).

Since c. d. b. must go into c. d. b. each of the following six σ's is possible. $\sigma_1(F_i) = f_i$, $\sigma_1(f_i) = F_i$ (as in Case II, interchanging A_i and B_i, where A_i and B_i are end points of F_i); $\sigma_2(F_1) = F_2$, $\sigma_2(F_2) = F_1$, $\sigma_2(f_1) = f_2$, $\sigma_2(f_2) = f_1$ (where A's go into B's and B's into A's); σ'_2 the same as σ_2, but

[14] Eilenberg, *loc. cit.*

[15] Since σ^2 leaves J pointwise fixed σ^2 is the identity.

A's go into A's and B's into B's; $\sigma_3(F_1) = f_2$, $\sigma_3(F_2) = f_1$, $\sigma_3(f_1) = F_2$, $\sigma_3(f_2) = F_1$ (A's into B's and B's into A's); σ'_3 same as σ_3 but A's go into A's and B's into B's; $\sigma_4(F_i) = F_i$, $\sigma_4(f_i) = f_i$ interchanging A_i and B_i. Now the following pairs cannot occur for the same mapping of G: (σ_1, σ_4), (σ_2, σ_3), (σ'_2, σ'_3). For suppose σ_1 and σ_4 both exist for a given G. Then $\sigma_4\sigma_1(F_i) = f_i$, $\sigma_4\sigma_1(f_i) = F_i$ leaving A_i and B_i fixed. But we have shown above (Case II) that there is no homeomorphism (other than the identity) leaving the end points of F_i fixed and carrying H_i into H_i. A similar contradiction may be obtained for the other two pairs of homeomorphisms. Therefore, there are at most three distinct homeomorphisms (excluding the identity) for any particular mapping of G.

The following relationships hold for the various σ's: $\sigma_2\sigma_1 = \sigma'_3$, $\sigma'_2\sigma_1 = \sigma_3$, $\sigma_2\sigma_4 = \sigma'_2$, $\sigma_3\sigma_4 = \sigma'_3$.

Furthermore, since each of the above σ's is such that σ^2 leaves at least one circuit of fixed points, it follows that $\sigma^2 = 1$ in each case (Eilenberg).

This establishes Case III, and the necessity of the given conditions.

The sufficiency of any one of the above conditions follows from Theorem C.

A further characterization of G when condition I is satisfied may be obtained by using Lemma 2. Let J be a split circuit with the invariant points X and Y. Then J can be expressed as the sum of two arcs, AXB and AYB, so that the feet of any component of $G - J$ lie entirely on AXB or entirely on AYB. Let N_1 be a component of $G - J$ with feet on AXB. Let $AXB = ApXqB$. Then for any point of $\bar{N}_1 \cdot J$ lying on $(ApX - X)$ there is a corresponding point of $\bar{N}_1 \cdot J$ lying on $(XqB - X)$.

In Case II, σ can be shown equivalent to a rotation of the sphere S through 180° about two invariant points, one lying in the region of S bounded by $f_1 + F_1$ that contains H_1, the other in the region of S bounded by $f_2 + F_2$ that contains H_2.

COROLLARY. *If G satisfies the conditions of Theorem D, then G either has no split circuits, or has four split circuits and at most four distinct homeomorphisms, or has at most two homeomorphisms.*

UNIVERSITY OF ARKANSAS,
AND
CORNELL UNIVERSITY.

GEOMETRIC CHARACTERIZATIONS OF INVARIANT PARTIAL DIFFERENTIAL EQUATIONS.*

By AARON FIALKOW.[1]

1. Introduction. This paper deals with a system of curves

$$f(x_1, x_2) = c \tag{1.1}$$

on an oriented two dimensional Riemannian surface element whose metric tensor is [2] g_{ij} so that

$$ds^2 = g_{ij}dx_i dx_j, \qquad |g_{ij}| > 0. \tag{1.2}$$

It is assumed that $f(x_1, x_2)$ is the solution of certain partial differential equations which involve the components of the metric tensor and their derivatives. The form of the equations which we consider is unchanged by an isometric transformation of the surface or by changes in the coördinate system (x_i). This means that the curves (1.1) which are solutions must have geometric peculiarities which we investigate.

For this purpose, it is useful to construct the orthogonal trajectories of (1.1). Then the geodesic curvatures of the two curves passing through each point and their tangential and normal derivatives are defined and are intrinsic quantities.[3] The intrinsic characterization of the system (1.1) which is subject to various invariant differential conditions is given as relations among these geodesic curvatures and their directional derivatives.[4] In the final sections, we derive equivalent geometric properties of a more synthetic nature if the surface is developable. We also indicate new characterizations of developable and minimal surfaces.

* Received December 14, 1936; revised May 5, 1937.

[1] National Research Council Fellow.

[2] As is customary in tensor analysis, an index which appears twice is to be summed over the range 1, 2.

[3] A detailed study of the intrinsic differential geometry on a surface appears in Graustein, "Méthodes Invariantes dans la Géométrie Infinitésimale des Surfaces," *Mémoires de l'Académie Royale de Belgique* (Classe des Sciences), (2), vol. 11 (1929), and "Invariant methods in classical differential geometry," *Bulletin of the American Mathematical Society*, vol. 36 (1930), pp. 489-521. Cf. in particular the sections on differential invariants.

[4] Similar characterizations are given by the author in *Proceedings of the National Academy of Sciences*, vol. 19 (1933), pp. 543-548.

2. Some invariant relations for an orthogonal net. A single infinitude of curves is given by (1.1). We may introduce new orthogonal curvilinear coördinates

$$u = u(x_1, x_2), \qquad v = v(x_1, x_2)$$

so that $v(x_1, x_2) = f(x_1, x_2)$. Then the given orthogonal net is equivalent to the parametric curves

$$v = c \tag{2.1}$$

$$u = c. \tag{2.2}$$

The metric tensor may be written as

$$ds^2 = e^{-2A(u,v)}du^2 + e^{-2B(u,v)}dv^2. \tag{2.3}$$

The geodesic curvature of a curve of (2.1) and its tangential and normal derivatives at a point are represented by γ, γ_s and γ_n. The symbols for the corresponding curve of (2.2) are Γ, Γ_S and Γ_N. An analogous notation is used for higher derivatives with the understanding that $\gamma_{ns} = \frac{d}{ds}\frac{d\gamma}{dn}$. The arc length of curves of (2.1) and (2.2) is denoted by s and S respectively. The positive tangents and normals are oriented so that

$$ds = -dN, \qquad dn = dS. \tag{2.4}$$

We shall use the following symbols for differential invariants which occur in our work:[5]

$$\Delta_1 f = g^{ij}f_{,i}f_{,j}$$
$$\Delta_2 f = g^{ij}f_{,ij}$$
$$\Delta_3 f = g^{ij}g^{lm}[f_{,i}f_{,j}f_{,lm} - f_{,i}f_{,l}f_{,jm}]$$
$$\Delta_4 f = \tfrac{1}{2}g^{ij}g^{lm}[f_{,i}f_{,jm} - f_{,ij}f_{,lm}].$$

Of course, $\Delta_1 f$ and $\Delta_2 f$ are the well known *first* and *second Beltrami parameters*. We note that since $|\, g_{ij}\,| > 0$, $\Delta_1 f > 0$. Hence division by $\Delta_1 f$ is always permissible—a fact we shall use frequently in what follows. When there is no ambiguity we shall write Δ_i for $\Delta_i f$.

In the coördinate system (u, v) having the metric tensor (2.3), the expressions for the geodesic curvatures of (2.1) and (2.2) as well as the Δ's is greatly simplified. This makes it easy to verify[6] that *the following relations hold for any orthogonal net:*

---

[5] The comma in the subscript denotes covariant differentiation as in Eisenhart, *Riemannian Geometry*, p. 26.

[6] Some of the details of the calculations necessary to derive (2.5) to (2.12) are

$$(\Delta_1)_s = -2\Gamma\Delta_1 \tag{2.5}$$

$$\Delta_2 = \tfrac{1}{2}\Delta_1^{-1/2}(\Delta_1)_n - \gamma\Delta_1^{1/2} \tag{2.6}$$

$$(\Delta_2)_s = -(\gamma_s + \Gamma_s)\Delta_1^{1/2} - 2\Gamma\Delta_2 \tag{2.7}$$

$$\Delta_3 = -\gamma\Delta_1^{3/2} \tag{2.8}$$

$$(\Delta_3)_s = (3\gamma\Gamma - \gamma_s)\Delta_1^{3/2} \tag{2.9}$$

$$\Delta_4 = \tfrac{1}{2}(\gamma(\Delta_1)_n - \Gamma(\Delta_1)_s). \tag{2.10}$$

The Gauss equation is equivalent to [7]

$$K = \gamma_n + \Gamma_N - \gamma^2 - \Gamma^2 \tag{2.11}$$

where K is the Gaussian curvature of the surface. Every point function $h(x_1, x_2)$ satisfies the integrability condition [8]

$$h_{sn} - h_{ns} = \gamma h_s + \Gamma h_n. \tag{2.12}$$

These equations are useful in the characterization of the solutions of various invariant types of differential equations.

We first prove that an analytic orthogonal net is completely determined by the values of γ, Γ and their derivatives at one point. By the classical theory, the values of

$$\gamma, \gamma_s, \gamma_{ss}, \cdots \tag{2.13}$$

and

$$\Gamma, \Gamma_S, \Gamma_{SS}, \cdots \tag{2.14}$$

determine the curves of (2.1) and (2.2) respectively passing through the point if the initial direction of one of them is given. We refer to these curves as the base curves. It will suffice to prove that through each point of the base curve of one family which is sufficiently near the given point there passes a unique curve of the orthogonal family. At any point of the base curve of (2.2), the corresponding curve of (2.1) depends on the values of the quantities (2.13) at this point. But these quantities are determined at any point of the base curve of (2.2) by equations of which

---

given in a forthcoming paper which characterizes the solutions of $\Delta_2 f = F(f)$ and $\Delta_2 f = af + b$ and gives some physical applications.

[7] Cf. Darboux, "Lecons Sur La Théorie Générale des Surfaces," vol. 3 (1894), p. 131.

[8] This condition is a special case of general integrability conditions which for an n-dimensional Riemannian space are due to Ricci. Cf. Eisenhart, *loc. cit.*, p. 99 and Graustein, *Bulletin, loc. cit.*, pp. 497-499.

$$\gamma = (\gamma)_0 + (\gamma_n)_0 \cdot S + \frac{1}{2!}(\gamma_{nn})_0 \cdot S^2 + \cdots$$

$$\gamma_s = (\gamma_s)_0 \cdot (\gamma_{sn})_0 \cdot S + \frac{1}{2!}(\gamma_{snn})_0 \cdot S^2 + \cdots$$

are the first two. Thus the corresponding curve of (2.1) is fixed. In the same way, we show that through each point of the base curve of (2.1), there passes a determinate curve of (2.2). This proves

THEOREM I. *The curves of an analytic orthogonal net are completely determined, except for initial direction, by the values of the geometric quantities* (2.13) *and* (2.14) *and their intrinsic normal derivatives of all orders at a single point.*

3. Geometric characterization of invariant orthogonal nets. In this section, we shall assume that $f(x_1, x_2)$ is the solution of an invariant partial differential equation and obtain characteristic properties of the corresponding orthogonal net. Such properties are known for some equations. Thus if $\Delta_1 f = F(f)$, then the curves (1.1) are parallel, i. e., $\Gamma = 0$. This well known result follows at once from (2.5) since $(\Delta_1)_s = F'(f) \cdot f_s$ and f_s is always identically zero. A characteristic property for the solutions of Laplace's equation, $\Delta_2 f = 0$, is also known.[9] It is $\gamma_s + \Gamma_s = 0$ and may easily be deduced from (2.7).

The two geometric equations given above involve the geodesic curvatures and their directional derivatives and are intrinsic quantities of the orthogonal net. On the other hand, the differential invariants although independent of the choice of the coördinate system (x_i) are not intrinsic since they depend upon the particular parameter f used to describe the curves (1.1). In what follows, we shall convert the non-intrinsic description of a system of curves by means of differential invariants into an intrinsic one by the use of geodesic curvatures. Valued suggestions by Dr. Harry Levy resulted in a simplification of our original proofs.

We now assume that

$$\Delta_4 f = F(f) \tag{3.1}$$

and discuss the geometry of the corresponding orthogonal net. From (3.1) it follows that $(\Delta_4)_s = 0$ since $f_s = 0$. We differentiate (2.10) with respect to s and set the right-hand member equal to zero. This gives

$$\gamma_s(\Delta_1)_n + \gamma(\Delta_1)_{ns} - \Gamma_s(\Delta_1)_s - \Gamma(\Delta_1)_{ss} = 0. \tag{3.2}$$

[9] Cf. Darboux, *loc. cit.*, vol. 3, p. 154.

From (2.12),

$$(\Delta_1)_{ns} = (\Delta_1)_{sn} - \gamma(\Delta_1)_s - \Gamma(\Delta_1)_n. \tag{3.3}$$

Substituting this expression in (3.2) and using (2.4) and (2.5) we obtain

$$(\Delta_1)_n[\gamma_s - 3\gamma\Gamma] - 2\Delta_1[\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3] = 0. \tag{3.4}$$

If

$$\gamma_s - 3\gamma\Gamma = 0 \tag{3.5}$$

then it follows that

$$\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3 = 0. \tag{3.6}$$

Conversely, if (3.5) and (3.6) are satisfied, (3.2) is true and (3.1) follows for some $F(f)$. Indeed it is easy to show that in this case if γ and Γ do not both vanish, if F is given, a suitable parameter $g(f)$ may be found so that

$$\Delta_4 g = F(g). \tag{3.7}$$

For $g(f) = c$ is the same system of curves as (1.1) and

$$\Delta_4 g(f) = g'(f)[g''(f) \cdot \Delta_3 f + g'(f) \cdot \Delta_4 f]. \tag{3.8}$$

From (3.5) and (2.9), $(\Delta_3 f)_s = 0$ or $\Delta_3 f = h(f)$. Since $\Delta_4 f = k(f)$ it follows from (3.8) that (3.7) has solutions $g(f)$ for all $F(g)$. The only exception occurs when $\Delta_3 f = \Delta_4 f = 0$. From (2.8) and (2.10) this means that $\gamma = \Gamma = 0$.

If $\gamma_s - 3\gamma\Gamma \neq 0$, we consider the system of equations consisting of (2.5) and (3.4). The condition of integrability of this system is (3.3). This becomes after simplifying by means of (2.5) and (3.4),

$$\begin{aligned}(\gamma_s - 3\gamma\Gamma)[&(\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3)_s + \Gamma(\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3) \\ &+ (\gamma_s - 3\gamma\Gamma)(\Gamma_S - \gamma\Gamma)] - (\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3)(\gamma_s - 3\gamma\Gamma)_s = 0.\end{aligned}$$

This proves

THEOREM II. *The necessary and sufficient condition that a one parameter family of curves on a surface, $f(x_1, x_2) = c$, be a solution of $\Delta_4 f = F(f)$ is that either $\gamma_s - 3\gamma\Gamma \neq 0$ and*

$$\begin{aligned}(\gamma_s - 3\gamma\Gamma)[&(\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3)_s \\ &+ \Gamma(\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3) + (\gamma_s - 3\gamma\Gamma)(\Gamma_S - \gamma\Gamma)] \\ &- (\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3)(\gamma_s - 3\gamma\Gamma)_s = 0 \qquad (\textit{Type I})\end{aligned}$$

or that

$$\gamma_s - 3\gamma\Gamma = 0$$

and $$\gamma\Gamma_S + 2\Gamma\Gamma_N - \gamma^2\Gamma + 2\Gamma^3 = 0. \qquad (Type\ II)$$

Any particular family of Type II which does not consist of parallel geodesics can be parametered so that it is a solution of $\Delta_4 f = F(f)$ *for every* $F(f)$.

By similar methods, the condition that $F(f)$ be constant may be obtained. By the last part of Theorem II, it is only necessary to consider families of Type I. In addition to the above property of Type I, we find another second order equation which guarantees that $F(f)$ is a constant. As this equation is complicated, we do not give it explicitly here. However, there is a simple characterization when

(3.9) $$\Delta_4 f = 0$$

which is derived below. These curves are related to developable surfaces in a manner which is considered in § 5.

If $\gamma \not\equiv 0$, from (3.9), (2.5) and (2.10) we have

(3.10) $$(\Delta_1)_n = -\frac{2\Gamma^2}{\gamma}\Delta_1.$$

This equation and (2.5) form a system whose integrability condition is (3.3). When this is simplified by means of (2.4), (2.5) and (3.10) we obtain

$$2\gamma\Gamma\Gamma_N + \Gamma^2\gamma_s + \gamma^2\Gamma_S - \gamma^3\Gamma - \gamma\Gamma^3 = 0.$$

It follows from (2.11) that this is equivalent to

$$\Gamma^2\gamma_s + \gamma\Gamma(\Gamma_N - \gamma_n) + \gamma^2\Gamma_S + \gamma\Gamma K = 0.$$

This result is stated in

THEOREM III. *The necessary and sufficient condition that a one-parameter family of curves on a surface,* $f(x_1, x_2) = c$, *which does not consist entirely of geodesics be a solution of* $\Delta_4 f = 0$ *is that*

$$\Gamma^2\gamma_s + \gamma\Gamma(\Gamma_N - \gamma_n) + \gamma^2\Gamma_S + \gamma\Gamma K = 0.$$

If $\gamma \equiv 0$, all the curves are geodesics. If they are solutions of (3.9), it follows from (2.5) and (2.10) that $\Gamma \equiv 0$, i. e., the geodesics are parallel. From (2.11), this can occur only on a developable surface. This shows that *the only systems of geodesics,* $f(x_1, x_2) = c$, *satisfying* $\Delta_4 f = 0$ *are families of parallel geodesics on a developable surface.*

More generally, if a one-parameter family of geodesics is a solution of (3.1), it is of Type II and according to Theorem II, $\Gamma(\Gamma^2 + \Gamma_N) = 0$. From

(2.11) and this equation we find that *the only systems of geodesics,* $f(x_1, x_2) = c$, *which are solutions of* $\Delta_4 f = F(f)$ *are such that* $\Gamma = \sqrt{-(K/2)}$. In particular, there cannot be a family of geodesics satisfying (3.1) on a region of a surface where the Gaussian curvature is positive.

We now determine the characteristic properties of the orthogonal net if

$$(\Delta_2 + \Delta_3) f = 0. \tag{3.11}$$

These curves are related to minimal surfaces as is shown in § 5. From (2.6) and (2.8) we find that (3.11) may be written as

$$(\sqrt{\Delta_1})_n = \gamma\sqrt{\Delta_1} + \gamma(\sqrt{\Delta_1})^3. \tag{3.12}$$

The identity (2.5) is equivalent to

$$(\sqrt{\Delta_1})_s = -\Gamma\sqrt{\Delta_1}. \tag{3.13}$$

The integrability condition of (3.12) and (3.13) is (2.12) where $h = \sqrt{\Delta_1}$. This condition, after simplification may be written as

$$(\gamma_s + \Gamma_S) + (\gamma_s - \gamma\Gamma)\Delta_1 = 0. \tag{3.14}$$

This means that the geometric characterization of (3.11) is obtained either by equating the coefficients of 1 and Δ_1 in (3.14) to zero or by substituting the value of Δ_1 given by (3.14) in both (3.12) and (3.13). In the first case the family is isothermal. This completes the proof of

THEOREM IV. *The necessary and sufficient conditions that a one-parameter family of curves on a surface,* $f(x_1, x_2) = c$, *be a solution of* $(\Delta_2 + \Delta_3) f = 0$ *is that either the family be non-isothermal and*

$$\left[\frac{\gamma_s + \Gamma_S}{\gamma_s - \gamma\Gamma}\right]_s + 2\Gamma\left[\frac{\gamma_s + \Gamma_S}{\gamma_s - \gamma\Gamma}\right] = 0$$

and

$$\left[\frac{\gamma_s + \Gamma_S}{\gamma_s - \gamma\Gamma}\right]_n - 2\gamma\left[\frac{\gamma_s + \Gamma_S}{\gamma_s - \gamma\Gamma}\right] + 2\gamma\left[\frac{\gamma_s + \Gamma_S}{\gamma_s - \gamma\Gamma}\right]^2 = 0$$

or that

$$\gamma_s + \Gamma_S = 0 \quad \textit{and} \quad \gamma_s - \gamma\Gamma = 0.$$

The geometric description of orthogonal nets which are solutions of $\Delta_3 f = F(f)$ follows immediately from (2.9) by setting the right-hand member equal to zero. Hence $\gamma_s - 3\gamma\Gamma = 0$ *is the characteristic property of the solutions of* $\Delta_3 f = F(f)$. Similarly, it follows from (2.8) that *any one-parameter system of geodesics is a solution of* $\Delta_3 f = 0$ *and conversely.*

4. Equivalent geometric properties on a developable surface. We consider some equivalent characterizations of the invariant systems of curves discussed in the previous section under the assumption that they are imbedded in a developable surface. In what follows, it is understood that the customary terminology of plane figures is to be applied to their isometric generalizations on any developable surface. The curves are referred to a system of coördinates (x, y) which is isometrically equivalent to plane Cartesian coördinates.

The characteristic property of curves which are solutions of $\Delta_3 f = F(f)$ as noted at the end of § 3 is

$$\gamma_s - 3\gamma\Gamma = 0. \tag{4.1}$$

The axis of deviation of a curve at a point is defined as follows: Chords are drawn parallel to the tangents at the point and terminating in the curve. The limiting position of the geodesic joining the midpoint of each chord to the point is the axis of deviation. Thus the axis is completely determined by the limiting position of four points which approach coincidence. Two of these determine the tangent; the other two fix a parallel chord. Hence the expression for the angle between the axis of deviation and the tangent involves only the first three derivatives. From analytic geometry, we find that if ϕ is the angle between the axis and the tangent geodesic

$$\tan \phi = -\frac{3\gamma^2}{\gamma_s}. \tag{4.2}$$

It is well known property of the parabola that the locus of the midpoints of parallel chords is a line parallel to the principal axis. Since the osculating parabola of a curve has third order contact with the curve, it follows that the axis of deviation at a point of a curve may also be defined as the geodesic through the point which is parallel to the principal axis of the corresponding osculating parabola.

The locus of points where the slope of the curves (1.1) is constant is called the system of *isoclines* or *isoclinal curves* of (1.1).[10] The equation of the isoclines is $f_x(x, y) + cf_y(x, y) = 0$ where c is an arbitrary constant. It is easy to show that the angle θ between a curve of (1.1) and the corresponding isocline is determined by

$$\tan \theta = -\frac{\gamma}{\Gamma}. \tag{4.3}$$

[10] The isogonal trajectories of a one-parameter family of curves consists of those curves which cut the family at a constant angle α. Each family obtained for a fixed value of α is called a base family of the system of isogonal trajectories. It is clear that the same isoclines are obtained for any base family. Thus the isoclines are a property of the whole isogonal system.

From (4.2) and (4.3) it follows that (4.1) is equivalent to $\phi = \theta$. This proves

THEOREM V. *The characteristic property of a one-parameter family of curves on a developable surfrace, $f(x_1, x_2) = c$, which is a solution of $\Delta_3 f = F(f)$, $(F \not\equiv 0)$ is that at each point the axis of deviation is tangent to the isoclinal curve.*

Straightforward calculation shows that the curvature of an isoclinal curve is

$$-\frac{\gamma^2\Gamma_S + \gamma\Gamma(\Gamma_N - \gamma_n) + \Gamma^2\gamma_s}{(\gamma^2 + \Gamma^2)^{3/2}}.$$

According to Theorem III, since $K = 0$, the characteristic property of curves which are not geodesics and which are solutions of $\Delta_4 f = 0$ is the vanishing of the numerator of this expression. Hence the property is equivalent to the condition that the isoclines be geodesics.[11] A family of curves of this kind is completely determined by any one-parameter family of geodesics and a single curve which cuts them. As a result, we have

THEOREM VI. *The only one-parameter families of curves on a developable surface, $f(x_1, x_2) = c$, which are solutions of $\Delta_4 f = 0$ are families of*

(1) *parallel geodesics*

(2) *curves whose isoclines are geodesics.*

If a system of curves is a solution of $\Delta_4 f = F(f)$ for every $F(f)$, it is of Type II. When $K = 0$, the second of the characteristic equations for this type given by Theorem II may be transformed by means of the first and (2.11) into

$$\Gamma^2\gamma_s + \gamma\Gamma(\Gamma_N - \gamma_n) + \gamma^2\Gamma_N = 0.$$

Hence by the results obtained in this section, the isoclines and the axes of deviation of the family must form the same system of geodesics. Since the isoclinal curves are geodesics, the differential equation of the family must be

$$y = x \cdot f(p) + g(p) \tag{4.4}$$

where $p = dy/dx$. Here $f(p)$ is the slope of an isocline on which the curves have slope p. Since the isocline is also the axis of deviation, the angle ϕ between the axis and the tangent geodesic is given by

[11] Since this condition also refers to the whole isogonal system, it follows that if one base family is a solution of $\Delta_4 f = 0$ then all are which do not consist entirely of geodesics.

$$\tan \phi = \frac{f(p) - p}{1 + p \cdot f(p)}. \tag{4.5}$$

A comparison of (4.2) and (4.5) leads to

$$\frac{y'''(1+p^2) - 3py''^2}{y''^2} = \frac{3(1 + p \cdot f(p))}{p - f(p)} \tag{4.6}$$

where γ and γ_s are represented by their equivalent expressions in the derivatives of y. Hence any family of Type II must be a common solution of (4.4) and (4.6).

After some discussion, we find that for the solutions of these two equations, either

$$f(p) = c; \qquad g(p) = \frac{a}{p} + b \tag{4.7}$$

or

$$f(p) = \frac{b - ap}{a + p}; \qquad g(p) = cf + d. \tag{4.8}$$

The solution of (4.4) and (4.7) is found to be any family of congruent parabolas with the same principal axis. The curves satisfying (4.4) and (4.8) form any system of similar ellipses or hyperbolas. That any family of similar ellipses or hyperbolas is of Type II follows at once, since it is a well known property of these curves that a diameter drawn to a point bisects the chords parallel to the tangent at that point. Of course, an analogous statement holds for the parabolas. This proves

THEOREM VII. *The only families of curves on a developable surface, $f(x_1, x_2) = c$, which are solutions of $\Delta_4 f = F(f)$ for every $F(f)$ are families of similar ellipses or hyperbolas and families generated by the translation of an arbitrary parabola parallel to its principal axis.*

5. Applications to the geometry of surfaces. Let an arbitrary surface be cut by all planes parallel to a fixed plane. If the single infinitude of plane sections is projected orthogonally upon the fixed plane, a one-parameter family of plane curves is obtained.[12] The following questions suggest themselves: What are the necessary geometric properties of this family if the surface is known to be of a given type defined by an invariant differential condition? Conversely, what properties of a one-parameter family of plane curves are

[12] We exclude the trivial case in which the family degenerates into a single curve. This occurs if a cylindrical surface is cut by planes perpendicular to its generators.

sufficient to guarantee a surface of this type such that the given curves are the projections of its parallel plane sections? The answers to these questions follow readily from our previous work in the case of developable and minimal surfaces. The proofs indicate how analogous results may be obtained for other types of surfaces.

Let

$$z = f(x, y) \tag{5.1}$$

be the equation of a surface which is cut by ∞^1 parallel planes

$$z = c. \tag{5.2}$$

Then the projections of the intersections of (5.2) with (5.1) are the family of plane curves (1.1). Since in the plane, $\Delta_4 f$ is the Hessian $f_{xy}^2 - f_{xx}f_{yy}$, the surface (5.1) is developable if and only if $\Delta_4 f = 0$. The conclusion of Theorem VI relative to this type of family enables us to characterize it geometrically. The necessary property of these curves may also be deduced directly from the character of the surface in a simple manner. A developable surface may be considered as generated by the tangent lines of a curve in space. The curve is the edge of regression of the surface. The tangent to a plane section at any point is determined by the intersection of the tangent plane of the surface with the plane of the section. But the tangent plane is the same at any point of a generator. Hence the angle between a fixed generator and the tangent to any plane section at a point where it meets this generator is constant. Therefore the projections of the generators are the isoclinal straight lines of the family of projected plane sections. The projection of the edge of regression is clearly the envelope of these straight lines.

In the degenerate case in which the family consists of straight lines, the plane sections on the surface must also be straight lines. Since the lines must be parallel the surface is a cylinder.

We note that if one function $f(x, y)$ is known which is a solution of $\Delta_4 f = 0$, the most general function is a solution of the equation obtained by setting the right-hand member of (3.8) equal to zero. If $\Delta_3 f \neq 0$, that is, if $\gamma \neq 0$, it follows that the most general function is $af + b$. This proves

THEOREM VIII. *The necessary and sufficient condition that a one-parameter family of plane curves that are not all straight lines be the projections of the ∞^1 parallel plane sections of a developable surface is that their isoclines be straight lines. The isoclines and their envelope are the projections of the*

generators and the edge of regression of the surface respectively. Any surface associated with a fixed family of curves in a fixed plane may be obtained from one such surface by a linear transformation orthogonal to the fixed plane. If the family of curves is composed entirely of straight lines, they must be parallel. In this case, the straight lines are the projections of the parallel generators of a cylindrical surface.

The surface (5. 1) is minimal if and only if $(\Delta_2 + \Delta_3)f = 0$ since in the plane, $(\Delta_2 + \Delta_3)f = (f_y^2 + 1)f_{xx} - 2f_xf_yf_{xy} + (f_x^2 + 1)f_{yy}$. Applying the results of Theorem IV, we obtain a characterization of minimal surfaces by means of the geometric properties of their parallel plane sections similar to that given above for developable surfaces.

PRINCETON UNIVERSITY
AND THE
INSTITUTE FOR ADVANCED STUDY.

SETS OF CONJUGATE MATRICES.*

By E. T. Browne.

1. Introduction. Let λ be an indeterminate scalar and let M be any given n-square matrix with elements independent of λ. If then $\phi(\lambda)$ is a scalar polynomial, and $M_1, M_2, \cdots, M_{\nu-1}$ are $\nu - 1$ matrices satisfying the conditions:

I. The matrices M_j $(j = 0, 1, \cdots, \nu - 1)$ are commutative in pairs;

II. The condition

$$(\lambda - M)(\lambda - M_1) \cdots (\lambda - M_{\nu-1}) \equiv \phi(\lambda)^1 \tag{1}$$

is satisfied identically in λ; we shall say that the $\nu - 1$ matrices $M_1, \cdots, M_{\nu-1}$ constitute a set of conjugates to M.

From (1) it is clear that the degree of the polynomial $\phi(\lambda)$ is ν. If ν is the *smallest* positive integer such that an identity of the type (1) subsists, we shall say that the $\nu - 1$ matrices M_j constitute a *reduced* set of conjugates to M. In the contrary case, we shall say that these matrices constitute an *extended* set of conjugates to M. From a reduced set of conjugates, one obvious way in which we can get an extended set is by adjoining to the former an arbitrary number of scalar matrices.

If in (1) we replace the scalar indeterminate λ by M, we see at once that $\phi(M) = 0$. Hence, $\phi(\lambda)$ must be divisible by the reduced characteristic function of M. However, it is known, and it will be clear from the results of this paper, that it suffices to take $\phi(\lambda)$ as the reduced characteristic function of M, so that henceforth in this paper $\phi(\lambda)$ will be used in precisely that sense.

The notion of a set of matrices conjugate to a given matrix was introduced by Taber,[2] who, however, confined his attention to matrices M of the third order with *distinct* characteristic roots. In this particular case, the reduced characteristic function coincides with the characteristic function $f(\lambda)$ of M. Taber imposed a third condition, *viz.*,

III. Each M_j has $f(\lambda)$ as its characteristic function.

In 1921 P. Franklin[3] employed the function $f(\lambda)$ on the right in (1)

* Received February 3, 1937.

[1] Here and elsewhere throughout this paper, the symbols λ and $\phi(\lambda)$ are employed to denote the scalar matrices λI, $\phi(\lambda) I$, where I denotes the unit matrix.

[2] Taber, *American Journal of Mathematics*, vol. 13 (1891), pp. 159-172.

[3] Franklin, *Annals of Mathematics*, 2nd Ser., vol. 23 (1923), pp. 97-100.

and showed that by omitting the condition III, a set of generalized conjugates could be set up for any square matrix M. E. S. Sokolnikoff [4] in 1933 was the first to employ the function $\phi(\lambda)$ in (1) and she was led thereby to a consideration of "matrices conjugate to a given matrix with respect to its minimum equation."

We also employ the reduced characteristic function $\phi(\lambda)$ on the right in (1) and it is the purpose of this paper to give an *à priori* method for deriving, along with other sets, the most general sets of conjugates which are expressible as polynomials in M. Before proceeding to the particular problem, however, we can make the following observation. If a set of $\nu - 1$ matrices exists satisfying I and II, and Sokolnikoff has shown that such sets do exist, we have on putting $\lambda = M_j$ in (1)

$$\phi(M_j) = 0, \qquad (j = 1, \cdots, \nu - 1).$$

We therefore have

THEOREM I.[5] *If $M_1, \cdots, M_{\nu-1}$ constitute a reduced set of conjugates to a matrix M, the reduced characteristic function of each M_j is a factor of the reduced characteristic function of M.*

Moreover, from the identity (1) we see at once that the elementary symmetric functions of the matrices $M, M_1, \cdots, M_{\nu-1}$ are equal to the elementary symmetric functions of the roots of $\phi(\lambda) = 0$, these latter functions being considered as scalar matrices.

2. The case in which M has a single characteristic root. We consider first the case in which M has a single characteristic root α, so that the reduced characteristic function of M reduces to

$$\phi(\lambda) = (\lambda - \alpha)^\nu \qquad (\nu \leqq n). \tag{2}$$

If we write

$$M = \alpha + \eta \tag{3}$$

i. e., $M - \alpha = \eta$, it is obvious from (2) that η is nilpotent of index ν, and that the matrices

$$1, \eta, \eta^2, \cdots, \eta^{\nu-1} \tag{4}$$

---

[4] Sokolnikoff, *American Journal of Mathematics*, vol. 55 (1933), pp. 167-180.

For a complete bibliography up to 1933 see MacDuffee, *The Theory of Matrices*, Berlin (Springer), 1933. Cf. also Hermann, "Über Matrixgleichungen und die Zerlegung von Polynomen in Linearfaktoren," *Compositio Mathematica*, vol. 1 (1934), pp. 284-302; Richardson, "Conjugate matrices," *Quarterly Journal of Mathematics*, vol. 7 (1936), pp. 256-270.

[5] Sokolnikoff, *loc. cit.*, p. 173.

are linearly independent. Moreover any polynomial in M can be written in the form

$$M_j = \alpha_j + \beta_j\eta + \gamma_j\eta^2 + \cdots + \kappa_j\eta^{\nu-1}, \tag{5}$$

and conversely, any matrix of the form M_j in (5) is a polynomial in M.

In this case (1) becomes

$$(\lambda - M)(\lambda - M_1)\cdots(\lambda - M_{\nu-1}) \equiv (\lambda - \alpha)^\nu, \tag{6}$$

and our first problem is to determine matrices M_j which are polynomials in M and which satisfy (6).

Since each M_j is expressible in the form (5), the left member of (6) is apparently a polynomial $F(\lambda, \eta)$ in λ and η. But the right member is free of η so that the left member must be also. Since λ^ν is the reduced characteristic function of η, and since in this connection η behaves in all ways precisely as a scalar, save that $\eta^\nu = 0$, this last condition is possible if, and only if, $F(\lambda, \eta)$ is divisible by η^ν; i. e., if and only if F and its first $\nu - 1$ derivatives as to η vanish for $\eta = 0$. Putting $\eta = 0$ in (6) we obtain as a first necessary condition

$$F(\lambda, 0) = (\lambda - \alpha)(\lambda - \alpha_1)\cdots(\lambda - \alpha_{\nu-1}) \equiv (\lambda - \alpha)^\nu,$$

identically in λ. From this we conclude that

$$\alpha = \alpha_1 = \alpha_2 = \cdots = \alpha_{\nu-1}. \tag{7}$$

Now designate by $M_j(\eta)$ the function on the right in (5) and differentiate $F(\lambda, \eta)$ partially as to η. We get

$$\frac{\partial F}{\partial \eta} = \sum_{j=0}^{\nu-1} \frac{\partial F}{\partial M_j} \cdot \frac{\partial M_j}{\partial \eta} = -\sum \frac{F(\lambda, \eta)}{\lambda - M_j} \cdot \frac{\partial M_j}{\partial \eta} = -(\lambda - \alpha)^\nu \sum \frac{\dfrac{\partial M_j}{\partial \eta}}{\lambda - M_j}.$$

Similarly,

$$\frac{\partial^2 F}{\partial \eta^2} = -(\lambda - \alpha)^\nu \sum \left\{ \frac{\dfrac{\partial^2 M_j}{\partial \eta^2}}{\lambda - M_j} + \frac{\left(\dfrac{\partial M_j}{\partial \eta}\right)^2}{(\lambda - M_j)^2} \right\};$$

$$\frac{\partial^3 F}{\partial \eta^3} = -(\lambda - \alpha)^\nu \sum \left\{ \frac{\dfrac{\partial^3 M_j}{\partial \eta^3}}{\lambda - M_j} + 3\frac{\dfrac{\partial M_j}{\partial \eta} \cdot \dfrac{\partial^2 M_j}{\partial \eta^2}}{(\lambda - M_j)^2} + \frac{2\left(\dfrac{\partial M_j}{\partial \eta}\right)^3}{(\lambda - M_j)^3} \right\};$$

and in general for $k = 4, 5, \cdots, \nu - 1$,

$$\frac{\partial^k F}{\partial \eta^k} = -(\lambda - \alpha)^\nu \sum \left\{ \frac{\dfrac{\partial^k M_j}{\partial \eta^k}}{\lambda - M_j} + \cdots + \frac{(k-1)!\left(\dfrac{\partial M_j}{\partial \eta}\right)^k}{(\lambda - M_j)^k} \right\},$$

where the dots indicate terms whose denominators are $(\lambda - M_j)^2, \cdots,$

$(\lambda - M_j)^{k-1}$. Since $(\lambda - \alpha)^\nu \not\equiv 0$, we have on putting $\eta = 0$, the following identities in λ:

$$(8) \qquad \Sigma \frac{\beta_j}{\lambda - \alpha_j} \equiv 0,$$

$$(9) \qquad \Sigma \left\{ \frac{2\gamma_j}{\lambda - \alpha_j} + \frac{\beta_j^2}{(\lambda - \alpha_j)^2} \right\} \equiv 0,$$

$$(10) \qquad \Sigma \left\{ \frac{6\delta_j}{\lambda - \alpha_j} + \frac{6\beta_j\gamma_j}{(\lambda - \alpha_j)^2} + \frac{2\beta_j^3}{(\lambda - \alpha_j)^3} \right\} \equiv 0,$$

$$(11) \qquad \Sigma \left\{ \frac{24\epsilon_j}{\lambda - \alpha_j} + \frac{24\beta_j\delta_j + 12\gamma_j^2}{(\lambda - \alpha_j)^2} + \frac{24\beta_j^2\gamma_j}{(\lambda - \alpha_j)^3} + \frac{6\beta_j^4}{(\lambda - \alpha_j)^4} \right\} \equiv 0,$$

. . . . . .

$$(12) \qquad \Sigma \left\{ \frac{\nu!\,\kappa_j}{\lambda - \alpha_j} + \cdots + \frac{(\nu-1)!\,\beta_j^{\nu-1}}{(\lambda - \alpha_j)^{\nu-1}} \right\} \equiv 0.$$

Since the α's are all equal, we conclude at once from (8) that

$$(13) \qquad \Sigma \beta_j = 0.$$

If we rewrite (9) in the form

$$2(\lambda - \alpha)\Sigma \gamma_j + \Sigma \beta_j^2 \equiv 0,$$

we see at once that also

$$(14) \qquad \Sigma \beta_j^2 = 0, \qquad \Sigma \gamma_j = 0;$$

and similarly from (10), (11) and (12)

$$(15) \qquad \Sigma \delta_j = \Sigma \epsilon_j = \cdots = \Sigma \kappa_j = 0,$$

$$(16) \qquad \Sigma \beta_j\gamma_j = \Sigma (2\beta_j\delta_j + \gamma_j^2) = \Sigma \beta_j^2\gamma_j = 0,$$

$$(17) \qquad \Sigma \beta_j^3 = \Sigma \beta_j^4 = \cdots = \Sigma \beta_j^{\nu-1} = 0.$$

Not only are these conditions necessary but they are sufficient that matrices M_j given by (5) shall satisfy the condition (6).

In view of (13), (14) and (17) it is obvious from Newton's identities that the β's are roots of an equation of the type

$$x^\nu - \beta^\nu = 0,$$

whence, since $\beta = 1$, the remaining β's are uniquely determined save for order. Indeed their values are

$$\omega, \omega^2, \cdots, \omega^{\nu-1}$$

where ω is a primitive ν-th root of unity.

The remaining conditions (14), $\cdots$, (16) are obviously satisfied if we take

$$\gamma_j = \delta_j = \epsilon_j = \cdots = \kappa_j = 0.$$

In this case we have as a set of conjugates to the matrix $M = \alpha + \eta$ the matrices

$$M_j = \alpha + \omega^j \eta \qquad (j = 1, \cdots, \nu - 1). \tag{18}$$

These conjugates were given first by P. Franklin, and were called by him *generalized* conjugates. They are obviously polynomials in M, and have not only the same characteristic function but the same elementary divisors as M.

However, there are other solutions than these for the γ's, δ's, etc. Indeed, since $\gamma = 0$, the ratios of the remaining γ's are determined uniquely from the $\nu - 2$ homogeneous linear equations

$$\begin{aligned} &\sum_1^{\nu-1} \gamma_j = 0, \\ &\sum \beta_j \gamma_j = 0, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ &\sum \beta_j^{\nu-3} \gamma_j = 0. \end{aligned} \tag{19}$$

Then, since the β's are distinct, the δ's are determined (but not uniquely) by the non-homogeneous equations

$$\begin{aligned} &\sum_1^{\nu-1} \delta_j = 0, \\ &2 \sum \beta_j \delta_j = -\sum \gamma_j^2, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \end{aligned} \tag{20}$$

etc.

For $\nu = 3$, we have as the most general set of conjugates which are expressible as polynomials in M

$$M = \alpha + \eta, \qquad M_1 = \alpha + \omega\eta + \gamma\eta^2, \qquad M_2 = \alpha + \omega^2\eta - \gamma\eta^2, \tag{21}$$

where ω is a complex cube root of unity and γ is arbitrary.

For $\nu = 4$, the β's are $1, i, -1, -i$, where $i = \sqrt{-1}$. Then from the relations

$$\begin{aligned} \gamma_1 + \gamma_2 + \gamma_3 &= 0, \\ i\gamma_1 - \gamma_2 - i\gamma_3 &= 0, \end{aligned}$$

the γ's are determined to be

$$\gamma_1 : \gamma_2 : \gamma_3 = 1 : i - 1 : -i,$$

while the δ's are subject to the single condition

$$\delta_1 + \delta_2 + \delta_3 = 0.$$

For $\nu = 4$, we have therefore the set of conjugates

$$\begin{aligned} M &= \alpha + \eta, \\ M_1 &= \alpha + i\eta + \rho\eta^2 + \delta_1\eta^3, \\ M_2 &= \alpha - \eta + \rho(i-1)\eta^2 + \delta_2\eta^3, \\ M_3 &= \alpha - i\eta - \rho i\eta^2 - (\delta_1 + \delta_2)\eta^3. \end{aligned} \tag{22}$$

We obtain in this way all reduced sets of conjugates to M which are expressible as polynomials in M. If $\nu = n$, i. e., if M has a single elementary divisor $(\lambda - \alpha)^n$, the only matrices commutative with M are polynomials in M. In this case the foregoing method yields *all* reduced sets of conjugates to M.

It should be noted particularly that the values of the β's, γ's, δ's, etc., depend in no wise on the value of α, but are dependent on ν only.

We now state a theorem which may be looked upon as essentially a corollary to Theorem I:

THEOREM II. *If M is a matrix of the form $\alpha + \eta$, where η is nilpotent of index ν, there does not exist a set of less than ν matrices, one of which is M itself, such that an identity of the type* (1) *subsists.*

3. Cyclic sets of conjugates. Referring to the matrix

$$M = \alpha + \eta$$

of the preceding section, we know that the matrix

$$M_1 = \alpha + \beta_1\eta + \gamma_1\eta^2 + \cdots + \kappa_1\eta^{\nu-1}$$

is a polynomial $\chi(M)$ in M. We inquire if it will be possible to determine the parameters $\gamma, \delta, \cdots$, etc., in such a way that the set of conjugates will be cyclic; i. e., so that

$$M_{j+1} = \chi(M_j) \qquad (j = 0, 1, \cdots, \nu - 1).$$

Here we require that $M_\nu = M$. Since in this connection η behaves in all ways precisely as a scalar, save that $\eta^\nu = 0$, we can expand $\chi(M)$ in ascending powers of η as follows

$$\chi(M) = \chi(\alpha + \eta) = \chi(\alpha) + \eta\chi'(\alpha) + \tfrac{1}{2}\eta^2\chi''(\alpha) + \cdots + \frac{1}{(\nu-1)!}\eta^{\nu-1}\chi^{(\nu-1)}(\alpha).$$

Since the matrices $1, \eta, \eta^2, \cdots, \eta^{\nu-1}$ are linearly independent the relation $M_1 = \chi(M)$ will be satisfied only if χ satisfies the conditions:

$$\chi(\alpha) = \alpha,\ \chi'(\alpha) = \beta_1,\ \chi''(\alpha) = 2\gamma_1, \cdots, \chi^{(\nu-1)}(\alpha) = (\nu-1)!\,\kappa_1.$$

Expanding $\chi(M_j) = \chi(\alpha + \beta_j\eta + \gamma_j\eta^2 + \cdots + \kappa_j\eta^{\nu-1})$ in an entirely similar manner, we obtain:

$$\chi(M_j) = \chi(\alpha) + \eta\beta_j\chi'(\alpha) + \frac{\eta^2}{2!}\{\beta_j^2\chi''(\alpha) + 2\gamma_j\chi'(\alpha)\} + \frac{\eta^3}{3!}\{\beta_j^3\chi'''(\alpha) + 6\beta_j\gamma_j\chi''(\alpha) + 6\delta_j\chi'(\alpha)\} + \cdots.$$

In view of the above conditions on χ, it follows that $M_{j+1} = \chi(M_j)$ if, and only if, the parameters β, γ, δ, etc., are determined so as to satisfy the conditions:[6]

$$(23)\quad \beta_{j+1} = \beta_j\chi'(\alpha) = \beta_j\beta_1,$$

$$(24)\quad \gamma_{j+1} = \tfrac{1}{2}\{\beta_j{}^2\chi''(\alpha) + 2\gamma_j\chi'(\alpha)\} = \beta_j{}^2\gamma_1 + \gamma_j\beta_1,$$

$$(25)\quad \delta_{j+1} = \tfrac{1}{6}\{\beta_j{}^3\chi'''(\alpha) + 6\beta_j\gamma_j\chi''(\alpha) + 6\delta_j\chi'(\alpha)\} = \beta_j{}^3\delta_1 + 2\beta_j\gamma_j\gamma_1 + \delta_j\beta_1,$$

$$(26)\quad \epsilon_{j+1} = \beta_j{}^4\epsilon_1 + 3\beta_j{}^2\gamma_j\delta_1 + (2\beta_j\delta_j + \gamma_j{}^2)\gamma_1 + \epsilon_j\beta_1,$$

etc. Here we recall that $\gamma = \delta = \epsilon = 0$ and we agree to interpret $\gamma_\nu = \gamma$, $\delta_\nu = \delta$, etc.

Since $\beta_1 = \omega$, $\beta_j = \omega^j$, the first of these conditions is always satisfied. The remaining conditions are also obviously satisfied in the case in which the γ's, δ's, etc., are taken to be zero.

We therefore have the theorem:

THEOREM III. *If M is a matrix with reduced characteristic function $(\lambda - \alpha)^\nu$ so that M can be written in the form $\alpha + \eta$ where η is nilpotent of index ν, then the $\nu - 1$ matrices*

$$M_j = \alpha + \omega^j\eta \qquad (j = 1, \cdots, \nu - 1)$$

form a cyclic set of conjugates to M; that is, there exists a polynomial χ such that

$$M_{j+1} = \chi(M_j) \qquad (j = 0, 1, \cdots, \nu - 1;\ M_\nu = M).$$

We inquire, however, if there does not exist a more general cyclic set of conjugates than this.

Suppose for example that $\nu = 3$ so that $\beta_1 = \omega$ is a cube root of unity. We have then only to consider the conditions (24) on the γ's, which reduce in this case to the single condition

$$\gamma_1 + \gamma_2 = 0.$$

It follows then that the set of conjugates in (21) is always cyclic, whatever value be assigned to γ. This may also be verified by direct calculation.

For the special case $\nu = 3$ we can therefore state the following result: *If M is a matrix with reduced characteristic function $(\lambda - \alpha)^3$, every reduced*

[6] These conditions could have been derived more easily as follows. Since $\eta = M - \alpha$, we have on substituting into the expression for M_1,

$$M_1 = \chi(M) = \alpha + \beta_1(M - \alpha) + \gamma_1(M - \alpha)^2 + \cdots + \kappa_1(M - \alpha)^{\nu-1}.$$

If now we replace M by M_j, M_1 by M_{j+1} expand the expression on the right and equate coefficients of corresponding powers of η we obtain precisely the conditions (23), ..., (26). However, for later purposes, it is more instructive to proceed as we did above.

set of conjugates to M, which are expressible as polynomials in M, is a cyclic set.

Next consider the case $\nu = 4$. The conditions (24) yield in this case precisely the values of the γ's that are given in (22), while (25) reduces merely to the following

$$\delta_2 = 2\rho^2 i, \qquad \delta_1 + \delta_2 + \delta_3 = 0.$$

Hence the set of conjugates in (22) becomes cyclic provided only δ_2 be taken equal to $2\rho^2 i$. This also may be verified by direct calculation from (22).

Thus we find that in the case $\nu = 4$, although not every set of conjugates which are polynomials in M will constitute a cyclic set, the parameters γ, δ can be assigned values, not all zero, such that the resulting set will be cyclic.

In equations (8), $\cdots$, (12) we have not given general relations connecting the parameters $\beta, \gamma, \cdots, \kappa$ for $\nu > 5$. The derivation of these depends on the general formula for the n-th derivative of a function of a function which is admittedly quite complicated. We shall, therefore, not attempt to give a general proof by the present method, but shall outline the proof for $\nu = 5$. The method seems to be general.

Consider first equation (24) connecting the γ's:

$$\gamma_{j+1} = \beta_1 \gamma_j + \gamma_1 \beta_j^2 \qquad (j = 0, 1, \cdots, 4). \tag{24}$$

Since $\gamma_0 = \gamma = 0$ by hypothesis, and since in order to complete the cycle we desire γ_5 to be equal to $\gamma = 0$, we have in (24) five homogeneous linear equations in the four unknowns $\gamma_1, \gamma_2, \gamma_3, \gamma_4$. The first of these reduces to the identity $\gamma_1 = \gamma_1$, so that there will be a non-zero solution for the four unknowns if, and only if, the four remaining equations are dependent. To show this dependence, multiply (24) through by β_1^{4-j} and sum as to j from 0 to 4. We get

$$\sum_0^4 \beta_1^{4-j} \gamma_{j+1} = \sum_0^4 \beta_1^{5-j} \gamma_j + \gamma_1 \sum \beta_1^{4+j}$$

since $\beta_j^2 = \beta_1^{2j}$. On the left change the index of summation by putting $j + 1 = i$. Since $\gamma_5 = \gamma_0 = 0$, this may be written:

$$\sum_0^4 \beta_1^{5-i} \gamma_i = \sum_0^4 \beta_1^{5-j} \gamma_j + \gamma_1 \beta_1^4 \sum_0^4 \beta_1^j.$$

Since β_1 is a primitive fifth root of unity, the last term on the right is zero and the remaining terms cancel. The equations (24) are therefore dependent. Indeed, $\gamma_2, \gamma_3, \gamma_4$ are uniquely determined in terms of γ_1.

We need now show that these γ's satisfy also the necessary conditions (19). To do this multiply (24) through by β_j^k $(k = 0, 1, 2)$, and sum as to j. We get

$$\sum_0^4 \beta_j{}^k \gamma_{j+1} = \beta_1 \sum \beta_j{}^k \gamma_j + \gamma_1 \sum \beta_j{}^{k+2}.$$

Since $\Sigma \beta_j{}^m = 0$ $(m = 2, 3, 4)$ the last summation on the right is zero. Also $\beta_j{}^k = \beta^k{}_{j+1}/\beta_1{}^k$, so that the above equation can be written, after changing as before the index of summation:

$$(1 - \beta_1{}^{k+1}) \sum \beta_j{}^k \gamma_j = 0 \qquad (k = 0, 1, 2).$$

Since $1 - \beta_1{}^{k+1} \neq 0$, we conclude that the γ's satisfying (24) satisfy (19) also.

In a similar manner we can show that there exist δ's and ϵ's not all zero which satisfy (25) and (26), and in addition satisfy (15) and (16). In fact by squaring (24) and summing it follows that for $\nu > 4$ the γ's satisfying (24) satisfy also the condition $\Sigma \gamma_j{}^2 = 0$, so that (16) reduces merely to $\Sigma \beta_i \delta_i = 0$.

As a very special solution we may take the γ's and δ's all zero and then we have only to determine the ϵ's satisfying (26). Thus, if we write $\beta_1 = \omega$, where ω is a primitive fifth root of unity, the conditions on the ϵ's become

$$\begin{aligned} \epsilon_2 &= (\omega + \omega^4)\epsilon_1 \\ \epsilon_3 &= \omega\epsilon_2 + \omega^3\epsilon_1 \\ \epsilon_4 &= \omega\epsilon_3 + \omega^2\epsilon_1 \\ 0 = \epsilon_5 &= \omega\epsilon_4 + \omega\epsilon_1. \end{aligned}$$

These equations yield the unique set of solutions

$$\epsilon_2 = (\omega + \omega^4)\epsilon_1, \qquad \epsilon_3 = -(\omega + \omega^4)\epsilon_1, \qquad \epsilon_4 = -\epsilon_1,$$

which obviously satisfy the additional condition $\Sigma \epsilon_j = 0$. Hence, in the case $\nu = 5$, we have the special set of conjugates:

$$\begin{aligned} M &= \alpha + \eta \\ M_1 &= \alpha + \omega\eta + \epsilon\eta^4 \\ M_2 &= \alpha + \omega^2\eta + (\omega + \omega^4)\epsilon\eta^4 \\ M_3 &= \alpha + \omega^3\eta - (\omega + \omega^4)\epsilon\eta^4 \\ M_4 &= \alpha + \omega^4\eta - \epsilon\eta^4. \end{aligned}$$

Here $M_{j+1} = \chi(M_j)$, where $\chi(\lambda) = \alpha + \omega(\lambda - \alpha) + \epsilon(\lambda - \alpha)^4$.

Moreover, in general, if $M = \alpha + \eta$, where η is nilpotent of index ν, and if we choose κ's such that

$$\begin{aligned} \kappa_0 &= 0, \\ \kappa_{j+1} &= \omega\kappa_j + \omega^{\nu-j}\kappa_1 \qquad (j = 0, 1, \cdots, \nu - 1) \end{aligned}$$

where ω is a primitive ν-th root of unity, it will follow that

$$\kappa_\nu = \kappa_0 = 0 \quad \text{and} \quad \sum_1^{\nu-1} \kappa_j = 0$$

and that the matrices

$$M_j = \alpha + \omega^j\eta + \kappa_j\eta^{\nu-1} \qquad (j = 1, \cdots, \nu-1) \tag{27}$$

form a cyclic set of conjugates to M. Indeed, it is easy to verify that if $\chi(\lambda)$ is the polynomial

$$\chi(\lambda) = \alpha + \omega(\lambda - \alpha) + \kappa_1(\lambda - \alpha)^{\nu-1}$$

then

$$M_{j+1} = \chi(M_j) \qquad (j = 0, \cdots, \nu-1).$$

We therefore have the theorem:

THEOREM IV. *If $M = \alpha + \eta$, where η is nilpotent of index ν, the matrices*

$$M_j = \alpha + \beta_j\eta + \gamma_j\eta^2 + \cdots + \kappa_j\eta^{\nu-1} \qquad (j = 1, \cdots, \nu-1)$$

are polynomials in M, and the parameters $\beta, \gamma, \cdots, \kappa$ can always be so chosen, with the γ's, $\cdots$, κ's not all zero, such the M_j constitute a cyclic set of conjugates to M.

It will be observed that for $\kappa_1 = 0$, the set of conjugates (27) reduces to the set of Theorem III.

4. Extended sets of conjugates. Hitherto we have considered reduced sets of conjugates only, and we have shown that if $M = \alpha + \eta$, where η is nilpotent of index ν, we can determine all sets of matrices $M_1, \cdots, M_{\nu-1}$ which are polynomials in M and which possess the property that

$$\prod_{j=0}^{\nu-1}(\lambda - M_j) = (\lambda - \alpha)^\nu. \tag{6}$$

If now η is of index $\mu < \nu$, the argument by which we obtained the conjugates in the preceding section shows that precisely the same coefficients that were found in determining the M_j there will yield a set of $\nu - 1$ M's satisfying (6). With regard to the matrix M, this set will in this case be not a reduced set but an extended set of conjugates.

For example, if

$$M = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix} = \alpha + \eta \quad \text{where} \quad \eta = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is nilpotent of index 2, the matrices in (21) with $\eta^2 = 0$

$$M = \alpha + \eta, \qquad M_1 = \alpha + \omega\eta, \qquad M_2 = \alpha + \omega^2\eta$$

are such that the last two constitute an extended set of conjugates to M. In this particular case, the reduced set of conjugates would consist of the single matrix $M_1 = \alpha - \eta$.

Before proceeding to a discussion of extended sets of conjugates we consider the following problem:

Let M be a matrix of the form $\alpha + \eta$, where η is nilpotent of index ν, and let m be a positive integer $< \nu$. Our problem is to find $m < \nu$ matrices N_j $(j = 1, \cdots, m)$ which are expressible as polynomials in M and which satisfy identically:

$$\prod_{j=1}^{m} (\lambda - N_j) = (\lambda - \alpha)^m. \tag{28}$$

If $m = 1$, N_1 must be a scalar α. We suppose then that $m > 1$. We may write N_1 in the form

$$N_1 = \alpha + a_1\eta + a_2\eta^2 + \cdots + a_{\nu-1}\eta^{\nu-1} = \alpha + \zeta,$$

where

$$\zeta = a_1\eta + a_2\eta^2 + \cdots + a_{\nu-1}\eta^{\nu-1} \tag{29}$$

is nilpotent of index $\leqq \nu$. If $a_1 \neq 0$, ζ is nilpotent of index ν, and by Theorem II we cannot find a set of N's including N_1 which satisfies (28).

Let a_k be the first coefficient in (29) that is different from zero. Then ζ is nilpotent of index μ where μ is such that

$$(\mu - 1)k < \nu \leqq \mu k.$$

We must determine k so that $\mu \leqq m$. Since $\nu > m$, we may write

$$\begin{aligned} \nu &= qm - r \qquad (0 \leqq r < m) \\ \nu &= (q-1)m + r' \qquad (0 < r' \leqq m), \end{aligned}$$

where q, r and r' are integers with $q > 1$. If we take $k \leqq q - 1$, then

$$km \leqq (q-1)m < \nu \leqq \mu k,$$

so that $\mu > m$. But if we take $k \geqq q$, then

$$km \geqq qm \geqq \nu > k(\mu - 1),$$

so that $m \geqq \mu$.

Let us then take $N_1 = \alpha + \zeta$, where

$$\zeta = a_q\eta^q + \cdots + a_{\nu-1}\eta^{\nu-1}.$$

ζ is then nilpotent of index $\mu \leqq m$ and we can determine, as in section 2, $\mu - 1$ matrices $N_2, \cdots, N_\mu$ which are polynomials in ζ and therefore in M, and such that

$$\prod_{j=1}^{\mu} (\lambda - N_j) = (\lambda - \alpha)^\mu.$$

If $\mu < m$, we may now adjoin to the set of μ matrices N_j $m - \mu$ matrices $N_{\mu+1}, \cdots, N_m$ which are scalars or polynomials in M, determined in the same manner as above, and such that for the entire set of N's we have

$$\prod_{j=1}^{m}(\lambda - N_j) = (\lambda - \alpha)^m.$$

It is clear that these matrices N can be built up in a great variety of ways.

Next suppose that $m > \nu$ in (28). Then an extended set of conjugates to M can be built up by the method mentioned at the beginning of the section. Otherwise, we may build up a reduced set of conjugates and adjoin to these $m - \nu$ scalars or matrices N built up in the manner just indicated.

Finally, we suppose that M is of the same form as before, *viz.*, $\alpha_1 + \eta$, and we set ourselves the problem of determining matrices M_j which are polynomials in M and which possess the property that

$$(30) \qquad (\lambda - M)(\lambda - M_1)\cdots(\lambda - M_{\tau-1}) = F(\lambda),$$

where

$$F(\lambda) = (\lambda - \alpha_1)^{\nu_1}(\lambda - \alpha_2)^{\nu_2}\cdots(\lambda - \alpha_r)^{\nu_r}, \qquad (\Sigma\nu_i = \tau).$$

Since $F(\lambda)$ must be divisible by the reduced characteristic function $(\lambda - \alpha_1)^{\nu}$ of M, this last condition is impossible if $\nu_1 < \nu$. We suppose then that $\nu_1 \geqq \nu$. We then determine first a set of conjugates $M_1, \cdots, M_{\nu_1-1}$, to M, which is either a reduced set or an extended set according as $\nu_1 = \nu$ or $\nu_1 > \nu$. These satisfy the condition:

$$\Pi(\lambda - M_j) = (\lambda - \alpha_1)^{\nu_1}.$$

Next we use the root α_2 and take $M_{\nu_1} = \alpha_2 + \zeta$ where ζ is of the form (29) and build up a set of $\nu_2 - 1$ conjugates to M_{ν_1}. These satisfy the condition

$$\Pi(\lambda - M_j) = (\lambda - \alpha_2)^{\nu_2}.$$

We proceed in the same way with regard to each of the factors $(\lambda - \alpha_i)^{\nu_i}$ of F. The set of matrices obtained by adjoining all these subsets are obviously commutative in pairs and satisfy (30).

5. The general square matrix M. We now turn to a consideration of a general square matrix M whose reduced characteristic function we shall suppose to be

$$(31) \qquad \phi(\lambda) = (\lambda - \alpha_1)^{\nu_1}(\lambda - \alpha_2)^{\nu_2}\cdots(\lambda - \alpha_r)^{\nu_r} \qquad (\Sigma\nu_i = \tau).$$

where the α's are distinct. In order to apply the results of the foregoing work to this case, we must first consider the principal idempotent and nilpotent elements of M.

6. The principal idempotent and nilpotent elements of M.[7] We sup-

[7] Wedderburn, "Lectures on matrices," *American Mathematical Society Publications*, 1934, pp. 28-29.

pose that the reduced characteristic function of M is given by (31) and that $r > 1$. Define the polynomials $h_i(\lambda)$ as follows

$$h_i(\lambda) = \frac{\phi(\lambda)}{(\lambda - \alpha_i)^{\nu_i}} \qquad (i = 1, \cdots, r).$$

We can then determine two scalar polynomials $H_i(\lambda)$ and $G_i(\lambda)$ of degrees not exceeding $\nu_i - 1$ and $\tau - \nu_i - 1$, respectively, such that

$$H_i(\lambda)h_i(\lambda) + G_i(\lambda)(\lambda - \alpha_i)^{\nu_i} \equiv 1.$$

If we write

$$\phi_i(\lambda) = H_i(\lambda)h_i(\lambda) \qquad (i = 1, \cdots, r),$$

and for the matric polynomial $\phi_i(M)$ write merely ϕ_i, then ϕ_i is the principal idempotent element of M corresponding to the root α_i. These matrices ϕ_i satisfy the conditions

$$\phi_i{}^k = \phi_i \qquad \text{for any positive integer } k; \tag{32}$$

$$\phi_i\phi_j = 0 \qquad (i \neq j); \tag{33}$$

$$\sum_{i=1}^{r} \phi_i = I. \tag{34}$$

Moreover, these ϕ's are linearly independent and none is zero.

Let us now denote by η_i the matric polynomial in M defined by

$$\eta_i = \eta_i(M) = (M - \alpha_i)\phi_i \qquad (i = 1, \cdots, r). \tag{35}$$

It is easily shown that these matrices η_i satisfy the conditions

$$\eta_i{}^k \neq 0 \quad (k < \nu_i); \qquad \eta_i{}^{\nu_i} = 0, \tag{36}$$

$$\eta_i\phi_i = \eta_i = \phi_i\eta_i; \qquad \eta_i\eta_j = 0 \quad (i \neq j), \tag{37}$$

and moreover,

$$M = \Sigma(\alpha_i\phi_i + \eta_i) = \Sigma\phi_i(\alpha_i + \eta_i). \tag{38}$$

The matrix η_i is the principal nilpotent element of M corresponding to the root α_i.

From the above relations it is easy to establish the following lemma:

LEMMA. *If* $\psi(\lambda) = a + b\lambda + c\lambda^2 + \cdots + k\lambda^{\nu_i - 1}$ *is any scalar polynomial in* λ, *the relation* $\phi_i\psi(\eta_i) = 0$ *is satisfied only if* $\psi(\lambda) \equiv 0$, *i. e., only if* $a = b = \cdots = k = 0$.

7. Sets of conjugates for the general matrix *M*. We set ourselves first the problem of determining $\tau - 1$ matrices M_j $(j = 1, \cdots, \tau - 1)$ which are expressible as polynomials in M and which in addition satisfy the condition:

$$(39)\quad (\lambda - M)(\lambda - M_1)\cdots(\lambda - M_{\tau-1}) \equiv \phi(\lambda) = (\lambda - \alpha_1)^{\nu_1}\cdots(\lambda - \alpha_r)^{\nu_r},$$

identically in λ. Such a set of M's automatically satisfies the condition I of section 1.

Since M can be written in the form (38) and since the ϕ's obey the relations (32) and (33), it follows readily that any polynomial $f(M)$ in M can be written

$$f(M) = \Sigma\phi_i f(\alpha_i + \eta_i) = \Sigma\phi_i f_i(\eta_i).$$

Conversely, since the ϕ's and η's are themselves polynomials in M, any expression of this last type, where the f_i are polynomials, is a polynomial in M. Let

$$M_j = \Sigma\phi_i f_i^{(j)}(\eta_i).$$

Since $\lambda = \lambda I = \Sigma\lambda\phi_i$, we may write (39) in the form

$$\Sigma\phi_i \prod_{j=0}^{\tau-1} \{\lambda - f_i^{(j)}(\eta_i)\} \equiv \Sigma\phi(\lambda)\phi_i,$$

which in view of (33) leads to

$$\phi_i \Pi\{\lambda - f_i^{(j)}(\eta_i)\} \equiv \phi(\lambda)\phi_i \qquad (i = 1, \cdots, r).$$

By the lemma of the preceding section, these conditions will hold if, and only if, the f's are so determined that

$$(40)\quad \Pi\{\lambda - f_i^{(j)}(\eta_i)\} \equiv (\lambda - \alpha_1)^{\nu_1}\cdots(\lambda - \alpha_r)^{\nu_r} \qquad (i = 1, \cdots, r).$$

Consider first this last condition with $i = 1$. Since M can be written in the form (38), the first factor in the brackets on the left is $\lambda - (\alpha_1 + \eta_1)$, where η_1 is nilpotent of index ν_1. Our problem therefore reduces to finding a set $f_1^{(j)}(\eta_1)$ $(j = 1, \cdots, \tau - 1)$ which are polynomials in η_1 and which constitute an extended set of conjugates to $\alpha_1 + \eta_1$. This problem was solved completely in section 4. We arrange these in a column headed ϕ_1, the leading entry in the column being $\alpha_1 + \eta_1$.

We then proceed in the same way to determine a set of $\tau - 1$ conjugates $f_i^{(j)}(\eta_i)$ with respect to $\alpha_i + \eta_i$ for $i = 2, \cdots, r$ in turn. For each i, we arrange these in a column headed ϕ_i and with $\alpha_i + \eta_i$ as the leading entry. In this way we are led finally to an array of the type:

$$(41)\quad \begin{array}{ccccc} \phi_1 & \cdots & \phi_i & \cdots & \phi_r \\ \alpha_1 + \eta_1 & \cdots & \alpha_i + \eta_i & \cdots & \alpha_r + \eta_r \\ f_1^{(1)}(\eta_1) & \cdots & f_i^{(1)}(\eta_i) & \cdots & f_r^{(1)}(\eta_r) \\ f_1^{(2)}(\eta_1) & \cdots & f_i^{(2)}(\eta_i) & \cdots & f_r^{(2)}(\eta_r) \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ f_1^{(\tau-1)}(\eta_1) & \cdots & f_i^{(\tau-1)}(\eta_i) & \cdots & f_r^{(\tau-1)}(\eta_r). \end{array}$$

Now M is equal to the sum of the products of the ϕ's by the corresponding elements in the first row of the above array. In the same way, we may take as M_j $(j = 1, \cdots, \tau - 1)$ the sum of the products of the ϕ's by the corresponding elements in the $(j+1)$-st row. Clearly, in any column of the above any two rows after the first can be interchanged. Hence a set of $\tau - 1$ conjugates M_j can be built up in a great variety of ways.

As an illustration, let us consider a matrix M which has as its reduced characteristic function

$$\phi(\lambda) = (\lambda - 1)^3(\lambda + 1)^2(\lambda - 2).$$

Here $\alpha_1 = 1$, $\alpha_2 = -1$, $\alpha_3 = 2$, so that M can be written

$$M = \phi_1 + \eta_1 - \phi_2 + \eta_2 + 2\phi_3$$

where η_1 and η_2 are nilpotent of index 3 and 2, respectively. The array built up in the manner described above would be in this case:

$$(42) \qquad \begin{array}{ccc} \phi_1 & \phi_2 & \phi_3 \\ 1 + \eta_1 & -1 + \eta_2 & 2 \\ 1 + \omega\eta_1 + a\eta_1^2 & -1 - \eta_2 & 1 \\ 1 + \omega^2\eta_1 - a\eta_1^2 & 2 & 1 \\ -1 + b\eta_1^2 & 1 + c\eta_2 & 1 \\ -1 - b\eta_1^2 & 1 - c\eta_2 & -1 \\ 2 & 1 & -1 \end{array}$$

Using the elements in the successive rows as coefficients of the ϕ's, we would have as one set of conjugates to M the following:

$$\begin{array}{llll} M_1 = \phi_1(1 + \omega\eta_1 + a\eta_1^2) + \phi_2(-1 - n_2) + \phi_3; & & & \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot & & & \\ M_5 = 2\phi_1 & + \phi_2 & + \phi_3. & \end{array}$$

The set of conjugates given by Sokolnikoff for this particular case is obtained from the above on putting $a = b = c = 0$.

We now prove the theorem:

THEOREM V. *If for a matrix M there exists a completely cyclic set of conjugates, it is always possible to arrange the elements in the columns of the array* (41) *in such a way that each column will be cyclic.*

For let

$$M_j = \sum_{i=1}^{\tau} \phi_i f_i^{(j)}(\eta_i)$$
$$M_{j+1} = \sum \phi_i f_i^{(j+1)}(\eta_i),$$

and suppose that there exists a polynomial χ such that

$$M_{j+1} = \chi(M_j) \qquad (j = 1, \cdots, \tau - 1).$$

Since $\chi(M_j) = \Sigma\phi_i\chi[f_i{}^{(j)}(\eta_i)]$ this requires that

$$\Sigma\phi_i f_i{}^{(j+1)}(\eta_i) = \Sigma\phi_i\chi[f_i{}^{(j)}(\eta_i)] \qquad (j = 1, \cdots, \tau - 1),$$

and this, by (33) and the lemma of section 6, requires that

$$f_i{}^{(j+1)}(\eta_i) = \chi[f_i{}^{(j)}(\eta_i)] \qquad (i = 1, \cdots, r;\ j = 1, \cdots, \tau - 1).$$

The theorem is therefore proved.

Now a mere glance at the third column of the array (42) is sufficient to show that the elements in this column cannot be made cyclic by any rearrangement of the elements.

We therefore have the theorem:

THEOREM VI. *It is not possible to obtain a completely cyclic set of conjugates for the general matrix M.*

8. Special cases. Consider now the special case considered by Sokolnikoff in which the reduced characteristic function of M is of the type

$$\phi(\lambda) = (\lambda - \alpha_1)^\nu(\lambda - \alpha_2)^\nu \cdots (\lambda - \alpha_r)^\nu.$$

We may then write

$$M = \Sigma(\alpha_i\phi_i + \eta_i)$$

where each η_i is nilpotent of index ν. Let ω denote a primitive ν-th root of unity. Then as given by Sokolnikoff,

the first $\nu - 1$ conjugates are $\Sigma(\alpha_i\phi_i + \omega^j\eta_i)$ $\qquad (j = 1, \cdots, \nu - 1)$,

the next ν conjugates are $\Sigma\alpha_{i+1}\phi_i$ $\qquad (\alpha_{r+i} = \alpha_i)$

the next ν conjugates are $\Sigma\alpha_{i+2}\phi_i$

and so on, where the α's are cyclicly interchanged.

These may be generalized in two ways.

In the first place, instead of taking the second set of ν conjugates as above, we might take ν matrices exactly the same as $M, M_1, \cdots, M_{\nu-1}$ above in every respect save that α_i has been replaced by α_{i+1}, i. e., with the α's cyclicly interchanged. For the next ν conjugates we might take ν matrices of the same type with the α's cyclicly interchanged once more; and so on.

On the other hand, instead of taking as the first ν conjugates the matrices given above, we might take the following:

$$M_j = \Sigma\phi_i(\alpha_i + \beta_j\eta_i + \gamma_j\eta_i{}^2 + \cdots + \kappa_j\eta_i{}^{\nu-1}) \qquad (j = 1, \cdots, \nu - 1),$$

the β's, γ's, $\cdots$, κ's being so determined as to satisfy the conditions (23), $\cdots$, (26). Since $\beta_j = \omega^j$, these conjugates reduce to those mentioned earlier

in this section provided we put $\gamma_j = \delta_j = \cdots = \kappa_j = 0$. The next ν conjugates may then be taken as of the same form as the matrices $M, M_1, \cdots, M_{\nu-1}$ but with the α's cyclicly interchanged; and so on.

It is then easy to show that the matrices $M, M_1, \cdots, M_{\nu-1}$ constitute a cyclic subset among the conjugates to M. To prove this, we note that M_1 is known to be a polynomial $\chi(M)$ in M.[8] Then just as in section 3 it follows that χ must satisfy the conditions

$$(43)\quad \chi(\alpha_i) = \alpha_i,\ \chi'(\alpha_i) = \beta_1,\ \chi''(\alpha_i) = 2\gamma_1, \cdots, \chi^{(\nu-1)}(\alpha_i) = (\nu-1)!\,\kappa_1.$$

In precisely the same way it can be shown that the further requirement $M_{j+1} = \chi(M_j)$ imposes only the additional conditions

$$\beta_{j+1} = \beta_j \chi'(\alpha_i) = \beta_j \beta_1,$$
$$\gamma_{j+1} = \tfrac{1}{2}[\beta_j{}^2 \chi''(\alpha_i) + 2\gamma_j \chi'(\alpha_i)] = \beta_j{}^2 \gamma_1 + \gamma_j \beta_1,$$

etc. But the β's, γ's, etc. were chosen in advance so as to satisfy these very conditions. Hence the statement is proved.

Similarly, it follows that each succeeding set of ν conjugates constitute a cyclic subset.

Consider now the r matrices of the above set:

$$M = \Sigma(\alpha_i \phi_i + \eta_i),$$
$$M_\nu = \Sigma(\alpha_{i+1} \phi_i + \eta_i),$$
$$\cdots\cdots\cdots$$
$$M_{(r-1)\nu} = \Sigma(\alpha_{i+r-1} \phi_i + \eta_i),$$

where each is obtained from the preceding by a cyclic interchange on the α's. We know that M_ν is a polynomial in M. From the equation

$$M_\nu = \chi(M) = \Sigma \phi_i \chi(\alpha_i + \eta_i) = \Sigma \phi_i [\chi(\alpha_i) + \eta \chi'(\alpha_i) + \frac{n^2}{2} \chi''(\alpha_i) + \cdots]$$

it follows that

$$\chi(\alpha_i) = \alpha_{i+1},\ \chi'(\alpha_i) = 1,\ \chi''(\alpha_i) = \cdots = \chi^{(\nu-1)}(\alpha_i) = 0.$$

But this is precisely the condition that

$$M_{(j+1)\nu} = \chi(M_{j\nu}).$$

It is easy to see that for $\nu > 1$ the conjugates

$$M_1, M_2, \cdots, M_{r\nu-1}$$

cannot be chosen in such a way that the entire set, including M, will be cyclic.

---

[8] Since the degree of the reduced characteristic function of M is νr any polynomial in M can be expressed uniquely as an equivalent polynomial of degree at most $\nu r - 1$. If then χ be taken of degree $\leqq \nu r - 1$ the expression of $M_1 = \chi(M)$ is unique.

For if $M_1 = \chi(M)$, the polynomial χ is subject to the conditions (43), while the condition

$$M_\nu = \chi(M_{\nu-1})$$

would require that $\chi(\alpha_i) = \alpha_{i+1}$ which is incompatible with (43).

We therefore have the theorem:

THEOREM VII. *If the reduced characteristic function of a matrix M is of the form $\phi(\lambda) = (\lambda - \alpha_1)^\nu(\lambda - \alpha_2)^\nu \cdots (\lambda - \alpha_r)^\nu$, then for $\nu > 1$ there does not exist a set of conjugates M_j which are polynomials in M and such that the entire set, including M, will be cyclic. However, the conjugates can be so chosen that in the set there will exist cyclic subsets, such as*

$$(M, M_1, \cdots, M_{\nu-1}), \cdots, (M_{(r-1)\nu}, \cdots, M_{r\nu-1}), \quad (M, M_\nu, \cdots, M_{(r-1)\nu}).$$

9. Matrices with linear elementary divisors. The case in which $\nu = 1$ is of particular interest. The roots of the minimum equation are then all distinct and M has linear elementary divisors. The conclusions here might be drawn from those in section 8 as corollaries.[9] However, it is believed that the case is of sufficient interest to warrant separate treatment.

Pierce[10] considered the case in which the characteristic equation of M has all distinct roots, which is precisely the situation considered originally by Taber for $n = 3$. Obviously, the situation considered by Pierce is a special case of the one treated in this section.

If $\nu = 1$ each η_i is zero so that M can be written

$$M = \Sigma \alpha_i \phi_i \tag{44}$$

and for any polynomial $\chi(M)$ we have

$$\chi(M) = \Sigma \chi(\alpha_i) \cdot \phi_i. \tag{45}$$

The array (41) has in this case as elements in its first row the numbers $\alpha_1, \alpha_2, \cdots, \alpha_r$. The columns are then built up by filling in with the α's in such a way that in each column each α is used once and once only. Here each of the conjugates M_j will be a polynomial in M, but it will not be true in general either that each M_j has the same reduced characteristic function as M or that the set will be cyclic.

For example, if $r = 3$ the array (41) might be

ϕ_1	ϕ_2	ϕ_3
α_1	α_2	α_3
α_2	α_1	α_1
α_3	α_3	α_2

[9] Sokolnikoff, *loc. cit.*, p. 175.

[10] Pierce, *Bulletin of the American Mathematical Society*, vol. 36 (1936), pp. 262-264.

corresponding to which we should have the set of conjugates

$$M = \alpha_1\phi_1 + \alpha_2\phi_2 + \alpha_3\phi_3$$
$$M_1 = \alpha_2\phi_1 + \alpha_1\phi_2 + \alpha_1\phi_3$$
$$M_2 = \alpha_3\phi_1 + \alpha_3\phi_2 + \alpha_2\phi_3.$$

The reduced characteristic function of M_1 is $(\lambda - \alpha_2)(\lambda - \alpha_1)$, while that of M_2 is $(\lambda - \alpha_3)(\lambda - \alpha_2)$. Moreover, if $M_1 = \chi(M)$ it follows from (45) that

$$\chi(\alpha_1) = \alpha_2, \qquad \chi(\alpha_2) = \alpha_1, \qquad \chi(\alpha_3) = \alpha_1.$$

But then $M_2 \neq \chi(M_1)$.

It is possible, however, to arrange the α's in the columns of (41) in such a way that both of these conditions will be satisfied. Indeed all we have to do is to arrange them in such a way that no α occurs more than once in any row or column. One way to do this is to interchange the α's cyclicly. Thus, if M is given in (44), then we might take as M_j the matrices

$$M_j = \Sigma\alpha_{i+j}\phi_i \qquad (j = 1, \cdots, r-1),$$

where for $i + j$ we take the least positive residue (mod r). Obviously if $M_1 = \chi(M)$, so that $\chi(\alpha_i) = \alpha_{i+1}$, then also

$$M_{j+1} = \chi(M_j).$$

Moreover, each M_j built up in this way has the same *reduced* characteristic function as M. However, it should be noted that this is not necessarily true as to the *characteristic* function.

THEOREM VIII. *If a matrix M has linear elementary divisors so that it can be written in the form* (44), *then for $r > 2$ there always exists more than one set of conjugates to M which are expressible as polynomials in M. It will not be true in general that each of the conjugates will have the same reduced characteristic function as M or that the set of conjugates will be cyclic. It is always possible, however, to find a set of conjugates satisfying both of these conditions. One such set is that obtained from the expression* (44) *of M by repeated cyclic interchanges on the α's.*

If $r = n$, the characteristic function and the reduced characteristic function are identical. In this case we have the

COROLLARY. *If the characteristic roots $\alpha_1, \cdots, \alpha_n$ of a matrix M are distinct there exists a polynomial $\chi(\lambda)$ (which is unique if its degree be taken $\leqq n-1$) such that*

$$\alpha_{i+1} = \chi(\alpha_i) \qquad (i = 1, \cdots, n;\ \alpha_{n+1} = \alpha_1).$$

If then we define

$$M_1 = \chi(M), M_2 = \chi(M_1), \cdots, M_{n-1} = \chi(M_{n-2}),$$

the matrices M_j all have the same characteristic function as M and constitute a cyclic set of conjugates to M.

10. Williamson's matrix.[11] In particular, if the characteristic function of M is of the form

$$\lambda^n - k^n \qquad (k \neq 0) \tag{46}$$

so that the characteristic roots of M are

$$\alpha_1 = k, \alpha_2 = \omega k, \cdots, \alpha_n = \omega^{n-1} k,$$

where ω is a primitive n-th root of unity, then the unique polynomial $\chi(\lambda)$ of degree $\leqq n-1$ such that $\chi(\alpha_i) = \alpha_{i+1}$ is precisely $\chi(\lambda) = \omega$. This is the case considered by Williamson.

Williamson's matrix is the product $B = \Omega A$ of the diagonal matrix $\Omega = [1, \omega, \cdots, \omega^{n-1}]$ by the circulant

$$A = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdot \\ a_1 & a_2 & \cdots & a_0 \end{pmatrix}.$$

If we can show that the characteristic function of B is of the form (46) with $k \neq 0$, then the matrices

$$\omega B, \omega^2 B, \cdots, \omega^{n-1} B \tag{47}$$

constitute a set of conjugates to B.

Write $\omega_i = \omega^i$, so that

$$1, \omega_1, \omega_2, \cdots, \omega_{n-1}$$

are the n-th roots of unity. The matrix

$$P = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_1 & \cdots & \omega_{n-1} \\ 1 & \omega_1^2 & \cdots & \omega^2_{n-1} \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \omega_1^{n-1} & \cdots & \omega^{n-1}_{n-1} \end{pmatrix}$$

is non-singular and, as is well known,[12] if we denote by $\theta(\omega_i)$ the polynomial

---

[11] Williamson, *American Mathematical Monthly*, vol. 39 (1932), pp. 280-285. See also Udo Wegner, "The product of a circulant matrix and a special diagonal matrix," *American Mathematical Monthly*, vol. 40 (1933), pp. 23-25.

[12] Scott, *Theory of Determinants*, Cambridge (1880), p. 82.

$$\theta(\omega_i) = a_0 + a_1\omega_i + a_2\omega_i^2 + \cdots + a_{n-1}\omega_i^{n-1},$$

we have

$$AP = PN, \quad \text{or} \quad P^{-1}AP = N$$

where N is the diagonal matrix $[\theta(1), \theta(\omega_1), \cdots, \theta(\omega_{n-1})]$. Moreover, it is easy to verify that $\Omega P = PR$, where R is the matrix

$$R = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Hence, $P^{-1}BP = P^{-1}\Omega PP^{-1}AP = RN$, where

$$RN = \begin{pmatrix} 0 & 0 & \cdots & 0 & \theta(\omega_{n-1}) \\ \theta(1) & 0 & \cdots & 0 & 0 \\ 0 & \theta(\omega_1) & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & \theta(\omega_{n-2}) & 0 \end{pmatrix}.$$

The characteristic function of this last matrix, and therefore of $B = \Omega A$, is then easily seen to be

$$\lambda^n - \theta(1)\theta(\omega_1)\cdots\theta(\omega_{n-1}) = \lambda^n - k^n$$

where k^n is equal to the determinant of ΩA or of $(-1)^{n-1}A$. If then $|A| \neq 0$ it follows that the matrices (47) form a cyclic set of conjugates to B.

But if $|A| = 0$, the characteristic function of B reduces to λ^n so that B is nilpotent of index $\nu \leqq n$. It follows then that

$$(\lambda - B)(\lambda - \omega B)\cdots(\lambda - \omega^{n-1}B) \equiv \lambda^n - B^n \equiv \lambda^n$$

so that the matrices (47) constitute a set, although not necessarily a *reduced* set, of conjugates to B.

11. Sets of conjugates for a derogatory matrix. In the case in which M has a single elementary divisor $(\lambda - \alpha_i)^{\nu_i}$ corresponding to each root α_i of the reduced characteristic equation $\phi(\lambda) = 0$ in (31), then the latter coincides with the characteristic equation. M is then said to be *non-derogatory*, and the only matrices M_j commutative with M are polynomials in M. In this case, therefore, all sets of conjugates to M are found by the methods of the preceding sections.

If, however, for at least one root α_i of $\phi(\lambda) = 0$ there is more than one

elementary divisor, the foregoing statement is no longer true. We proceed now to an investigation of this case.

Suppose for example that, corresponding to the root α_1, M has the elementary divisors

$$(48) \qquad (\lambda - \alpha_1)^{\nu_1}, (\lambda - \alpha_1)^{\nu'_1}, \cdots, (\lambda - \alpha_1)^{\nu_1^{(t)}}.$$

Then it is known that corresponding to the set of elementary divisors (48) there exists a set of so-called *partial* idempotent elements

$$\psi_0, \psi_1, \cdots, \psi_t,$$

and a set of partial nilpotent elements

$$\zeta_0, \zeta_1, \cdots, \zeta_t$$

the latter being nilpotent of indices

$$\nu_1, \nu', \cdots, \nu_1^{(t)}$$

respectively.

The ψ's satisfy relations similar to those satisfied by the principal idempotent elements ϕ_t in (32), (33) and (34); *viz.*,

$$\psi_i^k = \psi_i; \qquad \psi_i\psi_j = 0 \quad (i \neq j) \qquad \sum_{i=0}^{t} \psi_i = \phi_1,$$

where ϕ_1 is the principal idempotent element corresponding to the root α_1. These ζ's satisfy relations similar to (35), (36) and (37) satisfied by the principal nilpotent elements, *viz.*,

$$\zeta_i^k \neq 0 \quad (k < \nu_1^{(t)}), \qquad \zeta_i^{\nu_1^{(t)}} = 0, \qquad \zeta_i\psi_i = \zeta_i = \psi_i\zeta_i,$$
$$\zeta_i\zeta_j = 0 \quad (i \neq j); \qquad \Sigma\zeta_i = \eta_1$$

where η_1 is the principal nilpotent element corresponding to the root α_1.

These partial elements are not in general polynomials in M. Moreover they are not unique. In fact, if to the matrix M we apply the transformation $T^{-1}MT$, where T is commutative with M, a particular set of partial elements is in general transformed into another set, and all such sets are obtained in this way.

It should be clear from the properties of the partial elements that any matrix expressible as a polynomial in a set of partial elements of M is commutative with M, and further that any two matrices both of which are expressible as polynomials in the same set of partial elements are commutative with each other.

We seek then conjugates M_j $(j = 1, \cdots, \tau - 1)$ which are expressible

as polynomials in an arbitrary set of partial elements of M. It will then follow, just as in the discussion of section 7, that we are led to the problem of finding polynomials f satisfying equations of the form (40), where now η_1 has been replaced by each of the ζ_i in turn. In other words, in the array (41) the column headed ϕ_1 is split up into $t+1$ columns headed $\psi_0, \psi_1, \cdots, \psi_t$, respectively, and with the same root α_1 in all the columns.

In the first of these subcolumns, headed ψ_0, the leading entry is $\alpha_1 + \zeta_0$, where ζ_0 is nilpotent of index ν_1. The remaining $\tau - 1$ places in the first column are then to be filled in with matrices which are expressible as polynomials in ζ_0, and which constitute an extended set of conjugates to $\alpha_1 + \zeta_0$. These matrices satisfy a relation of the type (40) with η_1 replaced by ζ_0. In the second of these sub-columns headed ψ_1 the leading entry is $\alpha_1 + \zeta_1$. The remaining entries are matrices constituting an extended set of conjugates to $\alpha_1 + \zeta_1$, and satisfying a relation of the type (40) with η_1 replaced by ζ_1. We proceed in this way through all the $t+1$ columns, using $\zeta_2, \zeta_3, \cdots, \zeta_t$ in turn.

If there is more than one elementary divisor corresponding to the root α_2, we split up the column headed ϕ_2 into partial columns and proceed in the same way. Finally, we obtain an array of the type (41) but containing just as many columns as M has elementary divisors.

As the conjugate M_j $(j = 1, \cdots, \tau - 1)$ we may then take the sum of the products of the ψ's by the corresponding elements in the $(j+1)$-st row.

For a different set of partial elements clearly we are led to a different set of conjugates. We obtain in this way all sets of conjugates that are expressible as polynomials in the partial elements of M.

Consider for example a matrix of order 8 whose elementary divisors are

$$(\lambda - 1)^3, \qquad (\lambda - 1)^2, \qquad (\lambda + 1)^2, \qquad (\lambda - 2).$$

The partial elements corresponding to the last two elementary divisors are actually principal elements, so that the last two columns in the array (42) remain the same as there. However, the first column splits into two, as follows:

ψ_0	ψ_1	ϕ_2	ϕ_3
$1 + \zeta_0$	$1 + \zeta_1$	$-1 + \eta_2$	2
$1 + \omega\zeta_0 + a\zeta_0^2$	$1 - \zeta_1$	$-1 - \eta_2$	1
$1 + \omega^2\zeta_0 - a\zeta_0^2$	1	2	1
$-1 + b\zeta_0^2$	$-1 + d\zeta_1$	$1 + c\eta_2$	1
$-1 - b\zeta_0^2$	$-1 - d\zeta_1$	$1 - c\eta_2$	-1
2	2	1	-1.

If then we write

$$M = (1+\zeta_0)\psi_0 + (1+\zeta_1)\psi_1 + (-1+\eta_2)\phi_2 + 2\phi_3,$$

then $$M_1 = (1+\omega\zeta_0 + a\zeta_0^2)\psi_0 + (1-\zeta_1)\psi_1 + (-1-\eta_2)\phi_2 + \phi_3,$$

$$M_2 = (1+\omega^2\zeta_0 - a\zeta_0^2)\psi_0 + \psi_1 + 2\phi_2 + \phi_3,$$

etc.

Thus in the case where M is derogatory we have shown how to obtain many reduced sets of conjugates to M, but they are such that the M_j are not expressible as polynomials in M. However, even this method does not yield all sets of conjugates to M, since there exist matrices commutative with M which are not expressible as polynomials in any set of partial elements of M.

For example if

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

it is easy to verify that the M's are commutative in pairs and moreover satisfy the condition I of section 1, with $\phi(\lambda)$ replaced by $(\lambda-1)^3$, the characteristic function of M. Hence, M_1 and M_2 constitute an extended set of conjugates to M, but neither is expressible as a polynomial in any set of partial elements of M. Such sets of conjugates we leave out of consideration in the present paper, since the methods employed do not lend themselves readily to a discussion of them. They are found most readily, perhaps, by a method given by P. Franklin in his paper "Algebraic matric equations," *Journal of Mathematics and Physics*, vol. 10 (1932), pp. 289-314.

THE UNIVERSITY OF NORTH CAROLINA.

ABELIAN FIELDS AND DUALITY OF ABELIAN GROUPS.*

By REINHOLD BAER.

It is a classical theorem that the cyclic extensions of a (commutative) field are exactly those which are generated by solving an equation $x^n - k = 0$, provided sufficiently many roots of unity are available. Generalizing this result, Witt [1] has recently characterized the finite abelian extensions by means of invariants. It is the aim of this note to show that Witt's theory may be extended—almost without any modification—to all those extensions whose group is abelian (finite or not) and that this may be done by the simple device of substituting Pontrjagin's duality theory [2] for the classical theory of characters of finite abelian groups.

Chapter I. Vector Groups and Their Characters.

Instead of introducing a topology in the considered Galois groups and character groups it seems to be more appropriate to describe their elements as vectors whose coördinates are elements in certain finite groups and are subject to some conditions. This chapter is devoted to an account of these concepts and in particular to the theory of duality between abelian groups without elements of infinite order and abelian vector groups. Definitions and theorems are adapted to the needs of the applications in the second chapter and topological ideas have been avoided.

1. The concept of vector group. Two objects are needed for the definition of a vector group: a set S of groups and an operation $h(X < Y)$ defined for certain pairs X, Y in S. The groups in S will be called the coördinate groups and the group Y in S will be said to cover X in S, whenever $h(X < Y)$ exists. S and h will be subject to certain conditions which will be enunciated in the course of this section and which will be denoted by (1. A) and so on.

* Presented to the American Mathematical Society, September 6, 1937. Received by the Editors, June 16, 1937.

[1] E. Witt, "Der Existenzsatz für abelsche Funktionenkörper," *Journal für die reine und angewandte Mathematik*, Bd. 173 (1935); E. Witt, "Zyklische Körper und Algebren der Characteristik p vom Grad p^n," *Journal für die reine und angewandte Mathematik*, Bd. 176 (1937), pp. 126-140.

[2] L. Pontrjagin, "Theory of topological commutative groups," *Annals of Mathematics*, vol. 35 (1934), pp. 361-388; E. R. van Kampen, "Locally bicompact abelian groups and their character groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.

(1. A) *$h(X < X)$ exists for every X in S and is the identity automorphism of the group X.*
(1. B) *If Y covers X, then $h(X < Y)$ is a homomorphism, mapping the whole group Y upon the whole group X.*
(1. C) *If Z covers Y and Y covers X, then Z covers X and*

$$h(X < Z) = h(Y < Z)h(X < Y).$$

(1. D) *To any pair of groups in S there exists a group in S which covers both.*
(1. E) *To every group X in S and every normal subgroup N of X there exists one and only one group W in S such that X covers W and the homomorphism $h(W < X)$ maps exactly the elements in N upon* 1.
(1. F) *Every group in S is finite.*

It is a consequence of (1. A) and (1. E) that $h(X < Y)$ is an isomorphism if, and only if, *$X = Y$ and h is the identity.*

LEMMA 1. 1. *Suppose that S and h satisfy* (1. A) *to* (1. F).

(a) *If the group H in S covers the groups K and L in S, and if every element in H which is mapped upon* 1 *by $h(K < H)$ is also mapped upon* 1 *by $h(L < H)$, then K covers L.*
(b) *If F is a finite, not vacuous subset of S, then there exists one and only one group G in S such that G covers every group X in F and such that* 1 *is the only element in G which is mapped upon* 1 *by every homomorphism $h(X < G)$ for X in F. If furthermore H in S covers every group in F, then H covers G and G may therefore be called the least common cover ($= l. c. c.$) of the groups in F.*
(c) *If R is a not vacuous subset of S, then there exists one and only one group E in S which is covered by every group X in F and which covers every group in S which is covered by every group in F. E may therefore be called the greatest common ground ($= g. c. g.$) of the groups in R.*

Proof. Suppose that H covers the groups K and L. Denote by K' those elements in H *which are mapped upon* 1 *by* $h(K < H)$ and by L' those elements in H which are mapped upon 1 by $h(L < H)$. Assume that $K' \leq L'$. Then $h(K < H)$ maps L' upon a normal subgroup L'' of K and there exists therefore by (1. E) a group M in S such that K covers M and $h(M < K)$ maps exactly the elements in L'' upon 1. It follows now from (1. C) that H covers M and that $h(M < H)$ maps exactly the elements in L' upon 1 and it follows from (1. E) that $L = M$, i. e. that K covers L.

Suppose now that F is a finite, not vacuous subset of S. Then it follows

by complete induction from (1. C) and (1. D) that there exists at least one group H in S which covers every group X in F. Denote by H' the set of all those elements in the common cover H of F which are mapped upon 1 by every $h(X < H)$ for X in F. There exists by (1. E) a group G in S such that H covers G and $h(G < H)$ maps exactly the elements in H' upon 1, since H' is by (1. B) a normal subgroup of H. Now it follows from (a) that G is a common cover of the groups in F and it follows from the choice of G that 1 is the only element in G which is mapped upon 1 by every $h(X < G)$ for X in F.

Suppose now that $G(i)$ is—for $i = 1, 2$—a group in S such that $G(i)$ covers every X in F and such that 1 is the only element in $G(i)$ which is mapped upon 1 by every $h(X < G(i))$ for X in F. As has been proved just now, there exists a group G in S such that G covers $G(1)$ and $G(2)$ and such that 1 is the only element in G which is mapped upon 1 by both $h(G(1) < G)$ and $h(G(2) < G)$. It follows from (1. C) that G covers every X in S and that $h(X < G) = h(G(i) < G)h(X < G(i))$ for X in F and $i = 1, 2$. Thus if w is an element of G which is mapped upon 1 by $h(G(j) < G)$ for $j = 1$ or 2, then w is mapped upon 1 by every $h(X < G)$ for X in F and consequently w is mapped upon 1 by $h(G(j \pm 1) < G)$ and this implies that $w = 1$. Thus it follows from (1. E) and (1. A) that $G(1) = G = G(2)$.

It follows now from the proof that every common cover of the groups in F covers their uniquely determined l. c. c.

Denote by T the set of all those groups in S which are covered by every group in the subset R of S. It follows from (1. E) and (1. F) that T is finite, and T is not vacuous, since every group in S covers by (1. E) the uniquely determined group in S which consists of the identity only. It is now an obvious consequence of (b) that the l. c. c. of T is the g. c. g. of F and this completes the proof of the lemma.

If S is any set of groups, then it is usual to define as a vector any single valued function v which maps every group X in S on a uniquely determined element v_X of X, its X-coördinate. Vectors are multiplied in multiplying their coördinates and thus the vectors over S form a group. But if in the set S a covering operation $h(X < Y)$ is defined which satisfies (1. A) to (1. F), then only a certain subgroup will be considered, namely the subgroup of all those vectors v whose coördinates v_X satisfy:

(1. V) *If X covers Y, then $v_Y = v_X{}^{h(Y<X)}$.*

The group of all the vectors of S which satisfy (1. V) will be called the *vector group $V(S, h)$, defined over S by means of the covering operation $h(X < Y)$.*

If in particular S is a finite set, then it follows from Lemma 1.1, (b) that $V(S,h)$ is essentially the l. c. c. of S, and that S is essentially the set of all the different factor groups of $V(S,h)$.

If the set S is infinite, but countable, then there exists a sequence $H(i)$ $(i=1,2,\cdots)$ of groups in S such that $H(i)\neq H(j)$ for $i\neq j$, $H(i)$ covers $H(j)$ for $j<i$ and such that every group X in S is covered by some $H(i)$. The vector group $V(S,h)$ consists essentially of the sequences $h(i)$ with $h(i)$ in $H(i)$ and $h(i)=h(i+1)^{h(H(i)<H(i+1))}$. The sequences which satisfy $h(i)=1$ form a normal subgroup $V(i)$ of $V(S,h)$ and $V/V(i)$ and $H(i)$ are isomorphic. These groups are therefore essentially the so-called Cantorian groups.[3]

In the following the set S need not be countable and two further assumptions will be needed which are a consequence of (1. A) to (1. F), if S is at most countable.

(1. G) *If x is an element in the group X in S, then there exists a vector in $V(S,h)$ whose X-coördinate is x.*

In mapping every vector of $V(S,h)$ upon its X-coördinate a homomorphism of $V(S,h)$ upon the whole group X is defined, provided (1. G) is satisfied. If $V(X)$ is the set of all vectors in $V(S,h)$ whose X-coördinate is 1, then $V(X)$ is a normal subgroup and X and $V(S,h)/V(X)$ are essentially the same. Furthermore S is essentially equal to the set of these factor groups and h to the homomorphism induced by $V(S,h)$.

(1. H) *The subset T of the set S is the whole set S if, and only if, T satisfies the following conditions:*

(i) *If X is contained in T, then every group, covered by X, is contained in T.*

(ii) *If X and Y are contained in T, then T contains a group which covers both X and Y.*

(iii) *The vector v in $V(S,h)$ is the 1-vector if, and only if, $v_X=1$ for every X in T.*

2. Subgroups of vector groups. It will be assumed in this section 2. that the set S of coördinate groups, the covering operation $h(X<Y)$ and the vector group $V(S,h)$ satisfy (1. A) to (1. H).

Notations. If U is a subgroup of $V(S,h)$ and X a group in S, then U_X is the set of all the X-coördinates of vectors in U, $S(U)$ is the set of all those

[3] D. van Dantzig, "Zur topologischen Algebra III," *Comp. Math.*, Bd. 3 (1936), pp. 408-426; H. Freudenthal, "Entwicklungen von Räumen und ihren Gruppen," *Comp. Math.* Bd. 4 (1937), pp. 145-234.

groups X in S which satisfy $U_X = 1$; $V(U)$ denotes the set of all the vectors v in $V(S,h)$ such that $v_X = 1$ for every X in $S(U)$ and $\bar{U}$ denotes the set of all the vectors v in $V(S,h)$ such that v_X is for every X in S an element of U_X.

$V(U)$ and $\bar{U}$ are subgroups of $V(S,h)$.
$S(V(U)) = S(U)$, $\bar{U}_X = U_X$ for every X in S.
$V(V(U)) = V(U)$, $\bar{\bar{U}} = \bar{U}$.

If U is a normal subgroup of $V(S,h)$, then it follows from (1. G) that every U_X is a normal subgroup of X, and if conversely every U_X is a normal subgroup of X, then $\bar{U}$ is a normal subgroup of $V(S,h)$.

(2. 1) *$U \leq \bar{U} \leq V(U)$ for every subgroup U of $V(S,h)$.*

This is a consequence of the identity: $S(U) = S(\bar{U}) = S(V(U))$.

(2. 2) *If U is a normal subgroup of $V(S,h)$, then $V(U) = \bar{U}$.*

Proof. There exists by (1. E) to every group X in S a uniquely determined group X' in S such that X covers X' and $h(X' < X)$ maps exactly the elements in U_X upon 1. If X in S covers Y in S, then it follows from (1. V) that U_X is mapped by $h(Y < X)$ exactly upon U_Y. This implies in particular that $U_{X'} = 1$ for every X in S, i. e. that every X' is an element of $S(U)$, and $S(U)$ is therefore by (1. A) exactly the set of all the groups X' for X in S. If v is a vector in $V(U)$, then $v_{X'} = 1$ for every X in S and now it follows from the choice of the group X' that v_X is an element of U_X for every X in S, i. e. that v is a vector in $\bar{U}$. Hence $V(U) \leq \bar{U}$ for normal subgroups U of $V(S,h)$ and consequently $V(U) = \bar{U}$ by (2. 1).

DEFINITION 2. 3. *The subgroup U of $V(S,h)$ is closed, if $U = V(U)$.*

Thus all subgroups $V(U)$ are closed subgroups and normal subgroups U are closed, if $U = \bar{U}$.

If G is any group whatsoever, then *denote by $C(G)$ its commutator group* and by G^m for a positive integer m the subset of those elements of G which mod $C(G)$ are m-th powers of elements of G. Then $C(G)$ and G^m are characteristic subgroups of G, $G/C(G)$ and G/G^m are abelian groups and the orders of the elements of G/G^m are finite and divisors of m.

(2. 4) *$C(V(S,h))_X = C(X)$; $(V(S,h)^m)_X = X^m$ for every X in S.*

3. Characters of groups without elements of infinite order.

Notation. E is the essentially uniquely determined group with the following properties:

(i) *Every finite subset of E is contained in a cyclic subgroup of E;*
(ii) *E contains elements of every finite order.*

This group E is abelian, contains for every positive integer n exactly one cyclic subgroup of order n and is isomorphic with the multiplicative group of the roots of unity in the field of complex numbers.

Suppose now that G is an abelian group without elements of infinite order. Then every homomorphism of G into a subgroup of E is called a *character of* G and the product fg of the characters f and g of G is defined by the equation: $x^{(fg)} = x^f x^g$ for every x in G. Thus the characters of G form an abelian group.

Since homomorphisms map elements of the finite order n upon elements whose order is a divisor of n, as a rule not the whole group E will be needed for the definition of the characters and it is obvious which part of E is really necessary.

Every character of G induces a character in every subgroup of G and conversely it may be proved in the usual way [4] that every character of every subgroup of G is induced by a character of the whole group G. It is a consequence of these facts, of the finiteness of the orders of the elements in G and of the condition (ii), that 1 *is the only element in G which is mapped upon* 1 *by every character of G.*

If U is any subgroup of G, then denote by $H(U)$ the group of all the characters of U. Let $S = S(G)$ be the set of all the groups $H(U)$ for finite subgroups U of G. If $U \leq U'$, then every character of U' induces a character of U and thus a homomorphism $h(H(U) < H(U'))$ of $H(U')$ upon the whole group $H(U)$ is defined. S and h satisfy obviously (1. A) to (1. C) and (1. F). (1. D) is a consequence of the fact that the join of two finite subgroups of G is a finite subgroup of G and (1. E) is a theorem in the classical theory of characters of finite abelian groups.[5]

If c is any character of G, then c induces in every finite subgroup U of G a character $c_{H(U)}$, its $H(U)$-coördinate. The characters of G are completely determined by its $H(U)$-coördinates, since G is the join of its finite subgroups. By the definition of the homomorphisms $h(H(U) < H(U'))$ it follows that the coördinates of the characters of G satisfy (1. V) and thus it may be proved that *the group of characters of G is exactly the vector group $V(S, h)$, defined over $S(G)$ by means of h.*

(1. G) is a consequence of the fact that every character of a subgroup of G is induced by a character of G and (1. H) follows from the remark: if U is any subgroup of G and w an element of G which is not contained in U, then

[4] Cp. footnote [2].
[5] E. Hecke, *Vorlesungen über die Theorie der algebraischen Zahlen*, 1923.

there exists a character of G which maps every element of U upon 1, but w upon an element $\neq 1$.

If the group G is not abelian, but all the elements of $G/C(G)$ are of finite order, then every character of G (= homomorphism of G into E) maps the elements of $C(G)$ upon 1 and induces therefore a character of $G/C(G)$. Conversely every character of $G/C(G)$ is induced by exactly one character of G. Thus the characters of G are essentially the same as the characters of $G/C(G)$ and it is obvious how to apply the above considerations upon the theory of characters of groups G such that the elements of $G/C(G)$ are of finite order.

4. Characters of vector groups. If G is any group, f a homomorphism of G, then denote by $(G;f)$ the set of all those elements in G which are mapped upon 1 by f. $(G;f)$ is a normal subgroup of G and f defines an isomorphism of $G/(G;f)$ upon G^f.

In this section it shall be assumed that the set S of groups, the covering operation $h(X < Y)$ and the vector group $V(S,h)$ defined over S by means of h satisfy the conditions (1. A) to (1. H).

LEMMA 4. 1. *Let f be a homomorphism of $V(S,h)$ into the group E, defined in* 3. *Then there exists a group N in S such that $v^f = 1$, whenever $v_N = 1$ if, and only if, $(V(S,h);f)$ is a closed subgroup of the vector group $V(S,h)$ and $V(S,h)^f$ is a finite (and therefore a cyclic) subgroup of E.*

Proof. Assume first that there exists a group N in S such that $v_N = 1$ implies $v^f = 1$. Then $v_N = w_N$ implies $v^f = w^f$ and it follows therefore from (1. G) that f induces a homomorphism of N into E. $N^f = V(S,h)^f$ and (1. F) imply now that $V(S,h)^f$ is a finite and therefore a cyclic subgroup of E. A vector v belongs to $(V(S,h);f)$ if, and only if, its N-coördinate v_N belongs to $(N;f)$. Let now X be any group in S. Then there exists by (1. D) a common cover X' of X and N in S. If $v_{X'} = 1$, then $v_N = 1$ and therefore $v^f = 1$. Thus f is also a homomorphism of X' and the vector v belongs to $(V(S,h);f)$ if, and only if, its X'-coördinate belongs to $(X';f)$. But since $(X';f) = (V(S,h);f)_{X'}$ and since X' covers X, it follows that v belongs to $(V(S,f);f)$ if, and only if, its X-coördinate belongs to $(V(S,h);f)_X$. Thus $(V(S,h);f) = \overline{(V(S,h);f)}$ and $(V(S,h);f)$ is therefore, as a normal subgroup of $V(S,h)$, by (2. 2) a closed subgroup of $V(S,h)$.

Assume now that $(V(S,h);f)$ is a closed subgroup of $V(S,h)$ and that $V(S,h)^f$ is a finite and therefore a cyclic subgroup of E. There exists a vector g in $V(S,h)$ such that g^f generates the finite cyclic group $V(S,h)^f$. If $g_X{}^i$ is for every X in S an element of $(V(S,h);f)_X$, then g^i is an element

of $(V(S,h);f)$ since $(V(S,h);f)$ is a closed subgroup of the vector group $V(S,h)$. If n is the order of g mod $(V(S,h);f)$ and $0<i<n$, then there exists therefore a group $N(i)$ in S such that $g^{i}{}_{N(i)}$ is not an element of $(V(S,h);f)_{N(i)}$. By Lemma 1.1 there exists the l. c. c. N of these groups $N(i)$ in S. From the choice of the groups $N(i)$ and N it follows that g^i is an element of $(V(S,h);f)$ if, and only if, $g_N{}^i$ is an element of $(V(S,h);f)_N$. Let now v be any vector in $V(S,h)$. Then there exists an integer i such that vg^i is an element of $(V(S,h);f)$. Consequently $(vg^i)_N=v_Ng_N{}^i$ is an element of $(V(S,h);f)_N$. Now $v_N=1$ implies that $g_N{}^i$ is an element of $(V(S,h);f)_N$ and therefore that g^i is an element of $(V(S,h);f)$. Thus $v_N=1$ implies that v itself is an element of $(V(S,h);f)$ and this completes the proof of the Lemma.

DEFINITION 4.2. *A character of the vector group $V(S,h)$ is a homomorphism f of $V(S,h)$ into E such that $(V(S,h);f)$ is closed and $V(S,h)^f$ is finite.*

COROLLARY 4.3. *To every character f of the vector group $V(S,h)$ there exists one and only one group N_f in S such that $v^f=1$ if, and only if, $v_{N_f}=1$.*

N_f is uniquely determined by either of the following two properties:

f induces an isomorphism in N_f.

The group X in S covers N_f if, and only if, $v_X=1$ implies $v^f=1$.

This Corollary follows essentially from Lemma 4.1 and (1. E).

By means of this Corollary it is rather simple to construct all the possible characters of the vector group $V(S,h)$. Just choose any cyclic group Z in S and an isomorphism f of Z into E. Z and f determine uniquely a character f of $V(S,h)$ by the equation $v^f=v_z{}^f$, and $Z=N_f$. Conversely every character of $V(S,h)$ leads to such a pair Z, f which is uniquely determined by the character. This implies in particular that the vector v is mapped upon 1 by every character of the vector group $V(S,h)$ if, and only if, all its coördinates are elements of the respective commutator groups. *Thus if $V(S,h)$ is abelian, then 1 is the only vector mapped upon 1 by every character of the vector group $V(S,h)$.*

5. The duality theorem. Suppose that the set S of groups, the covering operation $h(X<Y)$ in S and the vector group $V(S,h)$, defined over S by means of h, satisfy the conditions (1. A) to (1. H), that G is a group such that all the elements of $G/C(G)$ *are of finite order*, and that an operation $g\circ v$ is defined which obeys the following rules:

(5. A) *$g \circ v$ is for every g in G and for every v in $V(S,h)$ a uniquely determined element of the group E, defined in section 3.*

(5. B) $(gh) \circ v = (g \circ v)(h \circ v)$, $g \circ (vw) = (g \circ v)(g \circ w)$.

(5. C) *The group $(V(S,h); g)$ of the elements v in $V(S,h)$ which satisfy $g \circ v = 1$ is for every g in G a closed subgroup of the vector group $V(S,h)$.*

These rules imply that every v in $V(S,h)$ induces by $g \circ v$ a character of the group G, and consequently $c \circ v = 1$ for every c in $C(G)$. Thus every g in G induces a homomorphism of $V(S,h)$ into a finite subgroup of E and it follows now from (5. C) that every element g in G induces by $g \circ v$ a character of the vector group $V(S,h)$.

It is of importance to note that for the considerations of this section not the whole group E is needed, but only a subgroup of E which contains elements of order n for every positive integer n which is the order of an element in $G/C(G)$ or the order of an element in some group $X/C(X)$ for X in S.

THEOREM 5. 1. *The following four properties of the group G, the vector group $V(S,h)$ and the operation $g \circ v$ are equivalent:*

(a) *$g \circ v = 1$ for every g in G if, and only if, $v = 1$ and*

$g \circ v = 1$ for every v in $V(S,h)$ if, and only if, $g = 1$.

(b) *$g \circ v = 1$ for every g in G if, and only if, $v = 1$,*

G is abelian and every character of G is induced by some v in $V(S,h)$.

(c) *$V(S,h)$ is abelian, every character of the vector group $V(S,h)$ is induced by some element in G and*

$g \circ v = 1$ for every v in $V(S,h)$ if, and only if, $g = 1$.

(d) *G and $V(S,h)$ are each the character group of the other group (and therefore abelian).*

Remark. If G and (or) $V(S,h)$ are *finite* groups, then the above Theorem 5. 1 is a well known fact in the theory of characters of finite abelian groups [5] and the propositions (a) to (d) together with the finiteness-assumption imply that G and $V(S,h)$ are *isomorphic.* This remark shall be used in the course of the proof of the Theorem 5. 1.

Proof. A. If G is abelian and every character of G is induced by an element of $V(S,h)$, then

$$g \circ v = 1 \text{ for every } v \text{ if, and only if, } g = 1,$$

as follows from a remark in section **3.**

If $V(S, h)$ is abelian and every character of the vector group $V(S, h)$ is induced by some element in G, then

$$g \circ v = 1 \text{ for every } g \text{ if, and only if, } v = 1,$$

as follows from a remark at the end of section **4.**

These facts imply that (b) and (c) are both consequences of (d) and that (a) is a consequence of each of the propositions (b) and (c). Thus the Theorem 5. 1 will be proved as soon as it has been shown that (d) is a consequence of (a).

B. Assume that (a) is satisfied.

Clearly both G and $V(S, h)$ are abelian groups, since the elements in G and in $V(S, h)$ induce homomorphisms into an abelian group.

If g is any element of the group G, then there exists by (5. C) and Corollary 4. 3 a uniquely determined group N_g in S such that $g \circ v = 1$ if, and only if, $v_{N_g} = 1$. If F is any finite subgroup of G, then there exists by Lemma 1. 1 the l. c. c. N_F of the groups N_g for g in F. Since N_F covers each N_g with g in F, and since 1 is the only element of N_F which is mapped upon 1 by every $h(N_g < N_F)$ with g in F, it follows from the definition of N_g that $g \circ v = 1$ for every g in F if, and only if, $v_{N_F} = 1$, and N_F is uniquely determined by this property. Since therefore

$$g \circ v = g \circ w \text{ for every } g \text{ in } F \text{ if, and only if, } v_{N_F} = w_{N_F},$$

an operation $g \circ x$ is defined for the elements g in F and x in N_F by the equations

$$g \circ x = g \circ v \text{ for every } v \text{ in } V(S, h) \text{ such that } v_{N_F} = x$$

and this operation between F and N_F satisfies (5. A) and (5. B). From (a) and the definition of $g \circ x$ it follows that $g \circ x = 1$ for every x implies $g = 1$, and it follows from the definition of N_F and the definition of $g \circ x$ that $g \circ x = 1$ for every g in F implies $x = 1$. It is therefore a consequence of the theory of characters of finite abelian groups—as has been mentioned in the Remark—that F and N_F are isomorphic finite abelian groups and each the character group of the other group (under the operation $g \circ x$).

The fundamental step in the proof will be the proof of the following statement:

(5.2) *N_F defines a one-one-correspondence between the set of all the finite subgroups F of G and the whole set S.*

Suppose first that F and F' are two finite subgroups of G such that $N_F = N_{F'} = M$. Denote by F'' the subgroup of G which is generated by the elements in F and in F'. F'' is a finite subgroup of G, since G is abelian. Since $g \circ v = 1$ for every g in F if, and only if, $v_M = 1$, and since $g' \circ v = 1$ for every g' in F' if, and only if, $v_M = 1$, it follows that $g'' \circ v = 1$ for every g'' in F'' if, and only if, $v_M = 1$. Thus $N_{F''} = M$ and the groups F, F', F'' and M are isomorphic. Since F and F' are subgroups of the isomorphic finite group F'', it follows that $F = F'' = F'$ and N_F defines therefore a one-one-correspondence between the set of all finite subgroups F of G and a certain subset T of S.

It is an obvious consequence of the definition of N_F that N_F covers $N_{F'}$ if, and only if, $F' \leq F$. If F and F' are any two finite subgroups of G, F'' their join-group, then $N_{F''}$ is a common cover of N_F and $N_{F'}$. If U is any subgroup of a group N_F, then denote by F' the group of all those elements g in F which satisfy $g \circ u = 1$ for every u in U. Since F and N_F are each the character group of the other group, it follows from the classical theory of characters of finite abelian groups [5] that $N_{F'}$ is covered by N_F and that exactly the elements in U are mapped upon 1 by $h(N_{F'} < N_F)$.—If finally v is a vector such that $v_{N_F} = 1$ for every finite subgroup F of G, then $f \circ v = 1$ for every element f of G and it follows from (a) that $v = 1$. The set T satisfies therefore the conditions (i) to (iii) of (1. H) and this implies $T = S$, thus completing the proof of (5.2).

If $\boldsymbol{c}$ is any character of the vector group $V(S, h)$, then there exists by Corollary 4.3 a group $N_{\boldsymbol{c}}$ in S such that $v^{\boldsymbol{c}} = 1$ if, and only if, $v_{N_{\boldsymbol{c}}} = 1$. The character $\boldsymbol{c}$ of $V(S, h)$ induces therefore a character $\boldsymbol{c}$ of the cyclic group $N_{\boldsymbol{c}}$. There exists by (5.2) a uniquely determined finite subgroup F of G such that $N_F = N_{\boldsymbol{c}}$ and F and N_F are each the character group of the other group. F contains therefore an element g such that $g \circ x = x^{\boldsymbol{c}}$ for every element x in $N_{\boldsymbol{c}}$ and consequently $g \circ v = v^{\boldsymbol{c}}$ for every v in $V(S, h)$. Thus every character of the vector group $V(S, h)$ is induced by an element of G.

If $\boldsymbol{c}$ is any character of the group G, then $\boldsymbol{c}$ induces in every finite subgroup F of G a character $\boldsymbol{c}_F$. Since N_F is the character group of F, there exists an element c_F in N_F such that $g \circ c_F = g^{\boldsymbol{c}_F}$ for every g in F. These elements c_F satisfy (1. V) and they are therefore the coördinates of a vector v in $V(S, h)$. If (g) is the cyclic subgroup of G, generated by g, then $g \circ v = g \circ v_{N_{(g)}} = g^{\boldsymbol{c}_{(g)}} = g^{\boldsymbol{c}}$ and the character $\boldsymbol{c}$ of G is therefore induced by the element v of $V(S, h)$. This completes the proof of the Theorem 5.1.

If the group G, the vector group $V(S,h)$ and the operation $g \circ v$ do not satisfy the properties (a) to (d) of the Theorem 5.1, then at least the following statements may be derived from this theorem:

denote by G_0 the set of all those elements g in G such that $g \circ v = 1$ for every v in $V(S,h)$ and denote by V_0 the set of all vectors v, satisfying $g \circ v = 1$ for every g in G; then

$$C(G) \leq G_0, \qquad \overline{C(V(S,h))} \leq V_0.$$

The operation $X \circ Y$, defined by $G_0 g \circ V_0 v = g \circ v$ for the elements X of G/G_0 and the elements Y of $V(S,h)/V_0$ satisfies (5.A) to (5.C) and the conditions (a) to (d) of Theorem 5.1. Every character of G is induced by an element of $V(S,h)$ and every character of the vector group $V(S,h)$ is induced by an element of the group G if, and only if, $G_0 = C(G)$ and $V_0 = \overline{C(V(S,h))}$.

Chapter II. Characters of Galois Groups and Characterization of Abelian Extensions.

6. Representation of Galois groups as vector groups. This section is devoted to a representation of Galois groups as vector groups and to an exposition of those parts of Galois theory which are needed for this purpose.

Notations. If F is a field and $\boldsymbol{G}$ a group of automorphisms of F, then $J(\boldsymbol{G})$ is the subfield of all those elements f in F which satisfy $f = f^{\boldsymbol{g}}$ for every $\boldsymbol{g}$ in $\boldsymbol{G}$.—If K is a subfield of F, then an automorphism $\boldsymbol{g}$ of F is called a K-automorphism of F, if $s = s^{\boldsymbol{g}}$ for every s in K, and the group of all the K-automorphisms of F is denoted by $\boldsymbol{R}(K < F)$.

THEOREM 6.1. *$\boldsymbol{G} = \boldsymbol{R}(J(\boldsymbol{G}) < F)$ for every finite group $\boldsymbol{G}$ of automorphisms of F.*

A Proof of this theorem may be added, since this theorem is slightly more general than the theorems which are usually reproduced in the literature, and since the proof itself is simpler than the proofs of the classical case use to be.[6]

It is obvious that $\boldsymbol{G} \leq \boldsymbol{R}(J(\boldsymbol{G}) < F)$. Assume now that the theorem is wrong. Then there exists an automorphism $\boldsymbol{w}$ in $\boldsymbol{R}(J(\boldsymbol{G}) < F)$ which is not contained in $\boldsymbol{G}$. There exists therefore to every element $\boldsymbol{g}$ in $\boldsymbol{G}$ an element $v(\boldsymbol{g})$ in F such that $v(\boldsymbol{g})^{\boldsymbol{g}} \neq v(\boldsymbol{g})^{\boldsymbol{w}}$.

Let now correspond to every element $\boldsymbol{g}$ in $\boldsymbol{G}$ a new symbol $x(\boldsymbol{g})$. If H is any field, then denote by $H^* = H(\cdots x(\boldsymbol{g}) \cdots)$ the field of all rational func-

[6] The idea of the proof goes back to a paper of O. Bolza.

tions in the $x(\boldsymbol{g})$ with coefficients in H, and if $\boldsymbol{h}$ is an automorphism of H, then denote by $\boldsymbol{h}^*$ the uniquely determined automorphism of H^* which coincides in H with $\boldsymbol{h}$ and has the $x(\boldsymbol{g})$ as fixed elements.

Let $\boldsymbol{G}^*$ be the group of the automorphisms $\boldsymbol{g}^*$ of F^* for $\boldsymbol{g}$ in $\boldsymbol{G}$. Then $J(\boldsymbol{G}^*) = J(\boldsymbol{G})^*$. Since $\boldsymbol{G}$ is a finite group, it is possible to introduce the quantity $t = \sum_{\boldsymbol{g} \text{ in } \boldsymbol{G}} v(\boldsymbol{g})x(\boldsymbol{g})$ in F^* and the polynomial

$$f(x) = f(x; \cdots x(\boldsymbol{g}) \cdots) = \prod_{\boldsymbol{h} \text{ in } \boldsymbol{G}} (x - t^{\boldsymbol{h}^*}).$$

Since the application of an automorphism $\boldsymbol{g}^*$, for $\boldsymbol{g}$ in $\boldsymbol{G}$, upon $f(x)$ effects a permutation of the factors, defining $f(x)$, the coefficients of the polynomial $f(x)$ are elements in $J(\boldsymbol{G}^*)$ and the coefficients of the polynomial $f(x; \cdots x(\boldsymbol{g}) \cdots)$ are elements of $J(\boldsymbol{G})$. Since $\boldsymbol{w}$ belongs to $\boldsymbol{R}(J(\boldsymbol{G}) < F)$, it follows that $\boldsymbol{w}^*$ belongs to $\boldsymbol{R}(J(\boldsymbol{G}^*) < F^*)$ and consequently that $t^{\boldsymbol{w}^*}$ is a solution of $f(x) = 0$. Hence there exists an automorphism $\boldsymbol{z}$ in $\boldsymbol{G}$ such that $t^{\boldsymbol{z}^*} = t^{\boldsymbol{w}^*}$. But since $t^{\boldsymbol{h}^*} = \sum_{\boldsymbol{g} \text{ in } \boldsymbol{G}} v(\boldsymbol{g})^{\boldsymbol{h}} x(\boldsymbol{g})$, this implies that $v(\boldsymbol{z})^{\boldsymbol{z}} = v(\boldsymbol{z})^{\boldsymbol{w}}$ in contradiction to the choice of the elements $v(\boldsymbol{g})$ and this completes the proof of the theorem.

THEOREM 6.2. *Let U be a subfield of the field F. Then every element of F is mapped by U-automorphisms of F only upon a finite number of different elements in F and $U = J(\boldsymbol{R}(U < F))$ if, and only if, every element in F is a solution of an equation in U whose degree is exactly the number of its different solutions in F (i.e. if F is an algebraic, normal and separable extension of U).*

This may be proved by a slight change in the usual procedure.[7] The decisive step may be stated as a lemma for future reference.

LEMMA 6.3. *If the field F is an algebraic and normal extension of its subfield U, then every U-isomorphism of a field between U and F upon a field between U and F is induced by an automorphism of F.*[8]

Suppose now that the field F is an algebraic, normal and separable extension of its subfield K. A field B between K and F is *normal* with regard to K, if every K-automorphism of F maps the field B upon B, and B is *finite*

[7] R. Baer, "Abbildungseigenschaften algebraischer Erweiterungen," *Mathematische Zeitschrift*, Bd. 33 (1931), pp. 451-479.

[8] E. Steinitz, *Algebraische Theorie der Körper*. Neu herausgegeben und mit einem Anhang: Abriss der Galoisschen Theorie versehen von Reinhold Baer und Helmut Hasse, Berlin, 1930. Cp. Lemma 1, Lemma 1a and Lemma 2 on pp. 134-136.

with regard to K, if B may be generated by adjoining a finite number of elements to K.

Denote by $S(K < F)$ the set of all the fields B between K and F which are finite and normal with regard to K and denote by $\boldsymbol{S}(K < F)$ the set of all groups $\boldsymbol{R}(K < B)$ for B in $S(K < F)$. If B and B' are two fields in $S(K < F)$ and $B \leq B'$, then every K-automorphism of B' induces a K-automorphism of B and every K-automorphism of B is induced, by Lemma 6.3, by some automorphism of B'. Thus a homomorphism $\boldsymbol{h}(\boldsymbol{R}(K < B) < \boldsymbol{R}(K < B'))$ of $\boldsymbol{R}(K < B')$ upon the whole group $\boldsymbol{R}(K < B)$ is defined. The set $\boldsymbol{S}(K < F)$ and the covering operation $\boldsymbol{h}$ satisfy obviously (1.A) to (1.C) and (1.F). That they satisfy (1.D) is a consequence of the fact that the field B'', generated by the fields B and B' in $S(K < F)$, belongs to $S(K < F)$. If B is a field in $S(K < F)$ and $\boldsymbol{N}$ a normal subgroup of $\boldsymbol{R}(K < B)$, then $J(\boldsymbol{N})$ is a field between K and B which is normal and finite with regard to K. $\boldsymbol{N}$ is finite, since $\boldsymbol{R}(K < B)$ is finite. Hence $\boldsymbol{N} = \boldsymbol{R}(J(\boldsymbol{N}) < F)$ by Theorem 6.1 and $\boldsymbol{h}(\boldsymbol{R}(K < J(\boldsymbol{N})) < \boldsymbol{R}(K < B))$ maps exactly the elements in $\boldsymbol{N}$ upon 1. Thus it has been shown that (1.A) *to* (1.F) *are fulfilled by the set* $\boldsymbol{S}(K < F)$ *and the covering operation* $\boldsymbol{h}$.

Every K-automorphism $\boldsymbol{g}$ of F induces an automorphism $\boldsymbol{g}_{\boldsymbol{B}}$ in every field B in $S(K < F)$. $\boldsymbol{g}_{\boldsymbol{B}}$ belongs to $\boldsymbol{B} = \boldsymbol{R}(K < B)$ and these coördinates satisfy (1.V). $\boldsymbol{g}$ is uniquely determined by the $\boldsymbol{g}_{\boldsymbol{B}}$, since F is the join of the fields B in $S(K < B)$. If in every group $\boldsymbol{B}$ in $\boldsymbol{S}(K < F)$ an automorphism $\boldsymbol{g}_{\boldsymbol{B}}$ has been chosen in such a way that (1.V) is satisfied, then these $\boldsymbol{g}_{\boldsymbol{B}}$ determine exactly one K-automorphism of F whose coördinates they are and this proves that *the group* $\boldsymbol{R}(K < F)$ *of the* K*-automorphisms of* F *is exactly the vector group, defined over the set* $\boldsymbol{S}(K < F)$ *by means of the covering operation* $\boldsymbol{h}$.

(1.G) is satisfied in this vector group, since by Lemma 6.3 every K-automorphism of a field B in $S(K < F)$ is induced by an automorphism of F. Suppose now that $\boldsymbol{T}$ is a subset of the set $\boldsymbol{S}(K < F)$ which satisfies the conditions (i) to (iii) of (1.H). Denote by T the set of the corresponding fields in $S(K < F)$ and by W the join of the fields in T. Then W is a field between K and F such that the identity is the only W-automorphism of F. It follows therefore from Theorem 6.2 that $W = F$ and this implies that $T = S(K < F)$ and consequently that $\boldsymbol{T} = \boldsymbol{S}(K < F)$, i.e. that also (1.H) is satisfied. The result of these considerations may be stated as a theorem.

THEOREM 6.4. *If the field* F *is an algebraic, normal and separable extension of its subfield* K, *then the set* $\boldsymbol{S}(K < F)$, *the covering operation* $\boldsymbol{h}$ *and the vector group* $\boldsymbol{V}(\boldsymbol{S}, \boldsymbol{h})$, *defined over* $\boldsymbol{S}(K < F)$ *by means of* $\boldsymbol{h}$, *obey the postulates* (1.A) *to* (1.H) *and* $\boldsymbol{V}(\boldsymbol{S}(K < F), \boldsymbol{h}) = \boldsymbol{R}(K < F)$.

COROLLARY 6.5. *If $\boldsymbol{G}$ is a subgroup of the vector group $\boldsymbol{R}(K < F)$, then $\boldsymbol{R}(J(\boldsymbol{G}) < F) = \boldsymbol{V}(\boldsymbol{G})$—in the notation of section 2—and if $\boldsymbol{G}$ is a normal subgroup, then—by (2.2)—$\boldsymbol{R}(J(\boldsymbol{G}) < F) = \bar{\boldsymbol{G}}$.*

7. Realization of the character group of the Galois group. Let F be an algebraic, normal and separable extension of the field K. It will be possible to use the group $E(K)$ of the roots of unity in K for the definition of the characters of the vector group $\boldsymbol{R}(K < F)$ if, and only if, the following hypothesis is satisfied:

(7.R) *K contains n different n-th roots of unity, if there are elements of order n in a group $\boldsymbol{B}/\boldsymbol{C}(\boldsymbol{B})$ for $\boldsymbol{B}$ in the set $\boldsymbol{S}(K < F)$.*

NOTATION. *K^** is the multiplicative group of all the elements $\neq 0$ in K and *$G^*(F)$ is the multiplicative group of all those elements $\neq 0$ in F whose order mod K^* is a positive integer n such that K contains n different n-th roots of unity.*

THEOREM 7.1. *Let F be an algebraic, normal and separable extension of the field K and G^* a group between K^* and $G^*(F)$.*

(a) *The operation $X \circ \boldsymbol{g}$, defined for the elements X of the group G^*/K^* and the elements $\boldsymbol{g}$ of the vector group $\boldsymbol{R}(K < F)$ by the equation*

$$K^* u \circ \boldsymbol{g} = u^{1-\boldsymbol{g}}$$

obeys the postulates (5.A) to (5.C).

(b) *The following properties of the field F, the group G^* and the operation $X \circ \boldsymbol{g}$ are equivalent:*

(1) *F is the field, generated by G^*.*

(2) *$\boldsymbol{R}(K < F)$ is abelian, F satisfies (7.R) and $G^* = G^*(F)$.*

(3) *G^*/K^* and $\boldsymbol{R}(K < F)$ are—by means of the operation $X \circ \boldsymbol{g}$—each the character group of the other group.*

Proof. The definition of $X \circ \boldsymbol{v}$ is independent of the particular choice of the representative of the class X, since the elements $\boldsymbol{v}$ are K-automorphisms of F and consequently $k^{1-\boldsymbol{v}} = 1$ for k in K^*. If the order of the element $K^* g$ of G^*/K^* is n, then g is a solution of the equation $x^n - g^n = 0$ with coefficients in K. $K^* g \circ \boldsymbol{v}$ is therefore for every K-automorphism $\boldsymbol{v}$ of F an

n-th root of unity and is by the choice of G^* an element of K, i. e. an element of $E(K)$.

$$(XY) \circ \boldsymbol{v} = (X \circ \boldsymbol{v})(Y \circ \boldsymbol{v}),$$

since $\boldsymbol{v}$ is an automorphism of F, and

$$X \circ (\boldsymbol{vw}) = X^{1-\boldsymbol{vw}} = X^{1-\boldsymbol{w}} X^{\boldsymbol{w}-\boldsymbol{vw}} = (X \circ \boldsymbol{w})(X \circ \boldsymbol{v})^{\boldsymbol{w}} = (X \circ \boldsymbol{w})(X \circ \boldsymbol{v})$$

since $X \circ \boldsymbol{v}$ is an element of K and $\boldsymbol{w}$ a K-automorphism. Thus (5. A) and (5. B) are satisfied. That (5. C) is satisfied, is a consequence of Lemma 4. 1 and the fact that every X in G^*/K^* is contained in a field B in $S(K < F)$.

If $X \circ \boldsymbol{v} = 1$ for every $\boldsymbol{v}$ in $\boldsymbol{R}(K < F)$, then X belongs to $J(\boldsymbol{R}(K < F))$ and since this field equals K by Theorem 6. 2 it follows that $X = K^* = 1$, i. e.

(7. 2) $\qquad X \circ \boldsymbol{v} = 1$ for every $\boldsymbol{v}$ if, and only if, $X = 1$.

If (1) is satisfied and $\boldsymbol{v}$ is an element of $\boldsymbol{R}(K < F)$ such that $X \circ \boldsymbol{v} = 1$ for every X in G^*/K^*, then $f^{\boldsymbol{v}} = f$ for every f in F, i. e. $\boldsymbol{v} = \boldsymbol{1}$ and it follows from (7. 2) that (a) of Theorem 5. 1 is satisfied. Theorem 5. 1 may be applied, as follows from Theorem 7. 1, (a), and the choice of G^* and consequently (d) of Theorem 5. 1 holds true. Hence (3) is a consequence of (1).

If (3) is satisfied, denote by W the field, generated by G^*. This is a field between K and F. Let $\boldsymbol{v}$ be any W-automorphism of F. Then $X \circ \boldsymbol{v} = 1$ for every X in G^*/K^*, since these X are subsets of W. Since (3) holds true, also (d) of Theorem 5. 1 holds true and therefore (a) of Theorem 5. 1 is satisfied. Thus $\boldsymbol{v} = \boldsymbol{1}$ and it follows from Theorem 6. 2 that $W = F$, i. e. (1) is a consequence of (3).

Since (a) of Theorem 7. 1 has been proved for every group G^* between K^* and $G^*(F)$, it may be applied upon $G^*(F)$ itself. The same applies to (7. 2). Thus if (3) holds true, there exists to every X in $G^*(F)/K^*$ an X' in G^*/K^* such that $X \circ \boldsymbol{v} = X' \circ \boldsymbol{v}$ for every $\boldsymbol{v}$ in $\boldsymbol{R}(K < F)$, and since this implies $X = X'$, it follows that (3) implies the equality of G^* and $G^*(F)$, i. e. (2).

Assume now that c is a character of the vector group $\boldsymbol{R}(K < F)$. Then there exists by Corollary 4. 3 a uniquely determined group $\boldsymbol{N}_c$ in $\boldsymbol{S}(K < F)$ such that $\boldsymbol{v}^c = 1$ if, and only if, $\boldsymbol{v}_{\boldsymbol{N}_c} = \boldsymbol{1}$. Denote by B_c the field in $S(K < F)$ such that $\boldsymbol{N}_c = \boldsymbol{R}(K < B_c)$. Then c is a character of the group $\boldsymbol{N}_c$. There exists in B_c an element u such that $b = \sum_{\boldsymbol{h} \text{ in } \boldsymbol{N}c} \boldsymbol{h}^c u^{\boldsymbol{h}} \neq 0$—this follows e. g. from the existence of a normal basis [9] of B_c with regard to K—and, using the fact

[9] M. Deuring, "Galoissche Theorie und Darstellungstheorie," *Mathematische Annalen*, Bd. 107 (1932), pp. 140-144.

that by (7. R) all the elements $\boldsymbol{h}^c$ are contained in K, it may be verified that $b^{1-h} = \boldsymbol{h}^c$. If n is the order of $\boldsymbol{N}_c$, then $b^n = \prod_{h \text{ in } \boldsymbol{N}c} b^h \boldsymbol{h}^c = b^{\Sigma h} \prod \boldsymbol{h}^c$ is an element of K and K^*b is consequently an element of $G^*(F)/K^*$. Since $K^*b \circ \boldsymbol{v} = \boldsymbol{v}^c$ for every $\boldsymbol{v}$ in $\boldsymbol{R}(K < F)$, it follows that (3) is a consequence of (2) and this completes the proof of the Theorem.

COROLLARY 7. 3. *Let F be an algebraic, normal and separable extension of the field K, satisfying* (7. R).

(i) $\boldsymbol{R}(K < F)$ *is abelian if, and only if, F is the field, generated by $G^*(F)$.—If $\boldsymbol{R}(K < F)$ is abelian, then $G^*(F)/K^*$ and the vector group $\boldsymbol{R}(K < F)$ are each the character group of the other group.*

(ii) $J(\boldsymbol{C}(\boldsymbol{R}(K < F)))$ *is the field, generated by $G^*(F)$.*

(iii) *Denote by $G^*{}_m(F)$ the group of elements in $G^*(F)$ whose order mod K^* is a divisor of m and by F_m the field, generated by $G^*{}_m(F)$, then $G^*(F_m) = G^*{}_m(F)$ and $F_m = J(\boldsymbol{R}(K < F)^m)$.*

(i) is an obvious consequence of the Theorem 7. 1 and (ii) and (iii) follow in using Corollary 6. 5 and (2. 4).

8. Characteristic invariants of abelian extensions.

DEFINITION 8. 1. *$e(K < F)$ is for every algebraic, normal and separable extension F of the field K the set of all the orders of elements in groups $\boldsymbol{B}/\boldsymbol{C}(\boldsymbol{B})$ for $\boldsymbol{B}$ in $\boldsymbol{S}(K < F)$ and*

$P(n, K < F)$ is for every positive integer n the group of all the elements $\neq 0$ in K which are n-th powers of elements in F.

$P(n, K < F)$ contains the group K^{*n} of the n-th powers of elements $\neq 0$ in K. If K contains exactly n different n-th roots of unity, then $P(n, K < F) = G^*{}_n(F)^n$ and *the correspondence between the elements of $G^*{}_n(F)$ and their n-th powers induces an isomorphism between $G^*{}_n(F)/K^*$ and $P(n, K < F)/K^{*n}$.*

THEOREM 8. 2. *Suppose that H and L are algebraic, normal and separable extensions of the field K, that $\boldsymbol{R}(K < H)$ and $\boldsymbol{R}(K < L)$ are abelian and that* (7. R) *is satisfied by H and L. There exists a K-isomorphism between H and L if, and only if,*

(a) $e(K < H) = e(K < L) = e$;

(b) $P(n, K < H) = P(n, K < L) = P(n)$ *for every n in e.*

Proof. It is obvious that (a) and (b) are necessary conditions. It follows from Corollary 7. 3 that H is the field, generated by $G^*(H)$, and L is generated by $G^*(L)$. Since $G^*(H)$ is the join of the groups $G^*_n(H)$ for n in e and $G^*(L)$ is the join of the groups $G^*_n(L)$ for n in e, it follows from (a) and (b) that H as well as L is a field, generated from K by solving all the equations

$$x^n - k = 0 \text{ for } n \text{ in } e \text{ and } k \text{ in } P(n).$$

The existence of a K-isomorphism between H and L is now a consequence of a well known theorem in the theory of fields.[10]

COROLLARY 8. 3. *Suppose that H and L are algebraic, normal and separable extensions of the field K, that $\mathbf{R}(K < H)$ and $\mathbf{R}(K < L)$ are abelian, that (7. R) is satisfied by H and L, and that the positive integer n is the l. c. m. of the integers in $e(K < H)$. There exists a K-isomorphism between H and L if, and only if,*

(a′) *n is the l. c. m. of the integers in $e(K < L)$;*

(b′) $P(n, K < H) = P(n, K < L)$.

This is a consequence of the proof of Theorem 8. 2, since the conditions concerning n imply that $G^*(F) = G^*_n(F)$ for $F = H, L$.

Remark. If the fields H and L are contained in a field M, e. g. in the algebraic closure of K, then the conditions of Theorem 8. 2 or Corollary 8. 3 imply not only the existence of a K-isomorphism between H and L, but even the *equality* of H and L, since by Lemma 6. 3 every K-isomorphism of fields between K and its algebraic closure is induced by an automorphism of this algebraic closure, and since furthermore normal, algebraic extensions are mapped upon themselves by K-automorphisms of the algebraic closure of K.

THEOREM 8. 4. *Suppose that e is a set of positive integers, containing all the divisors of its elements and the l. c. m. of every finite subset, that K is a field, containing n different n-th roots of unity for every n in e, and that $D(n)$ is—for every n in e—a multiplicative subgroup of K^*.*

There exists an algebraic, normal and separable extension F of K such that $\mathbf{R}(K < F)$ is abelian, $e(K < F) = e$ and $P(n, K < F) = D(n)$ for every n in e if, and only if,

(1) $K^{*n} \leq D(n)$;

[10] E. Steinitz, *loc. cit.*, p. 112.

(2) *$D(n)/K^{*n}$ contains elements of order n;*

(3) $D(n)^k \leq D(nk)$;

(4) *the elements of $D(nk)$ whose order mod K^{*nk} divides n are represented exactly by the elements of $D(n)^k$.*

Proof. A. Suppose that F is an algebraic, normal and separable extension of the field K, that (7. R) is satisfied and that $\boldsymbol{R}(K < F)$ is abelian. Then $K^{*n} \leq P(n, K < F)$, since $K^* \leq G^*{}_n(F)$. $e(K < F)$ contains exactly the orders of the elements in the character group of the vector group $\boldsymbol{R}(K < F)$. $P(n, K < F)/K^{*n}$ contains therefore elements of order n, if n is contained in $e(K < F)$, since $P(n, K < F)/K^{*n}$ and $G^*{}_n(F)/K^*$ are isomorphic and $G^*(F)/K^*$ is the character group of $\boldsymbol{R}(K < F)$. If u is an element of $P(n, K < F)$, then there exists an element v in F such that $v^n = u$. $u^k = v^{nk}$ is therefore an element of $P(nk, K < F)$. If u is an element of $P(nk, K < F)$ whose order mod K^{*nk} is a divisor of n, then there exists an element f in F such that $f^{nk} = u$ and an element h in K^* such that $u^n = h^{nk}$. Consequently $(f^n)^{nk} = u^n = h^{nk}$ and $f^n h^{-1}$ is an nk-th root of unity. Since the nk-th roots of unity are by (7. R) elements of K, it follows that f^n is an element of K and u is therefore an element of $P(n, K < F)^k$. If conversely the element u in $P(nk, K < F)$ is at the same time an element of $P(n, K < F)^k$, then $u = v^k$ for a suitable element v in k^* and $u^n = v^{nk}$ is therefore an element of K^{*nk}. Thus it has been proved that the conditions (1) to (4) are necessary conditions.

B. Assume now that the groups $D(n)$ satisfy the conditions (1) to (4). Then denote by F the essentially uniquely determined field which is generated from K by completely solving the equations

$$x^n - d = 0 \text{ for } n \text{ in } e \text{ and } d \text{ in } D(n),$$

and denote by G^* the set of all the solutions of these equations in F.

Since K contains for every n in e exactly n different n-th roots of unity, these numbers n are relatively prime to the characteristic of K and F is therefore [11] an algebraic, normal and separable extension of K.

If $u(i)$ is for $i = 1, 2$ an element of G^*, then there exists a number $n(i)$ in e such that $u(i)^{n(i)}$ is an element of $D(n(i))$ and there exists the l. c. m. m of $n(1)$ and $n(2)$ in e. Since $m = n(i)k(i)$, it follows that $u(i)^m = (u(i)^{n(i)})^{k(i)}$ is an element of $D(n(i))^{k(i)}$ and therefore by (3) an element of $D(m)$. Since $D(m)$ is a group, $(u(1)u(2))^m$ is an element of

[11] E. Steinitz, *loc. cit.*, p. 66.

$D(m)$ and $u(1)u(2)$ is consequently an element of G^*. G^* is therefore by condition (1) and by the definition of G^* a group between K^* and $G^*(F)$. Since F is generated by G^*, it follows from Theorem 7. 1, (b) that

$$\boldsymbol{R}(K < F) \text{ is abelian, } F \text{ satisfies (7. R) and } G^* = G^*(F).$$

Since e is the set of orders of elements in G^*/K^*, it follows from Theorem 7. 1 that $e = e(K < F)$.

It is an obvious consequence of the definition of F and G^* that $D(n) \leq P(n, K < F)$ for every n in e. Assume now that u is any element in $P(n, K < F)$. Then there exists an element v in $G^*(F)$ such that $v^n = u$. Since $G^* = G^*(F)$, there exists an integer m in e such that v^m is an element of $D(m)$. Denote by q the l. c. m. of n and m. q is an element of e and $q = nn' = mm'$. Now it follows from condition (3) that $v^q = (v^m)^{m'}$ is an element of $D(q)$ and, since the order mod K^{*m} of $v^q = u^{n'}$ is a divisor of n, it follows from condition (4) that $u^{n'} = v^q = w^{n'}$ for some w in $D(n)$. uw^{-1} is therefore an n'-th root of unity and thus contained in K, since n' is a divisor of q. Since the n'-th roots of unity are for the same reason elements of K^{*n}, it follows from (1) that $u = (uw^{-1})w$ is an element of $D(n)$. Hence

$$P(n, K < F) = D(n) \text{ for every } n \text{ in } e,$$

and this completes the proof of the Theorem.

COROLLARY 8. 5. *Suppose that the field K contains n different n-th roots of unity and that P is a multiplicative group of elements $\neq 0$ in K. Then there exists an algebraic, normal and separable extension F of K such that $\boldsymbol{R}(K < F)$ is abelian, n is the l. c. m. of the numbers in $e(K < F)$ and $P(n, K < F) = P$ if, and only if $K^{*n} \leq P$ and P/K^{*n} contains elements of order n.*

Remark. In this exposition of a theory of abelian extensions it has been assumed that the set $e(K < F)$ does not contain the characteristic of the field K. If the characteristic of the field K is a prime number p and $e(K < F)$ consists of powers of p only, then it is again possible to extend the methods developed by Witt for the finite case to the general case. But a more detailed discussion of this case may be omitted, since it may be handled in exactly the same way and the translation from the case "prime to characteristic" to the other case follows exactly the lines, indicated by Witt.

THE UNIVERSITY OF NORTH CAROLINA,
CHAPEL HILL, N. C.

A NOTE ON TOPOLOGICAL FIELDS.*

By N. JACOBSON.[1]

In a recent paper [2] the author showed that a locally compact, separable, totally disconnected (l. c. s. t. d.) field $\mathfrak{F}$ has a valuation determined as follows: Let $\mathfrak{J}$ be the totality of elements a, called integers, such that $\{a^\nu\}$ is bounded (contained in a compact and closed set)[3] and $\mathfrak{P}$ the subset of elements b such that $b^\nu \to 0$. Then $\mathfrak{J}$ is a compact and open domain of integrity and $\mathfrak{P}$ a two-sided principal prime ideal $= (x)$ in $\mathfrak{J}$. Every element y in $\mathfrak{F}$ has the form ux^k where k is a rational integer and $u \in \mathfrak{J} \mid \mathfrak{P}$ the part of $\mathfrak{J}$ not in $\mathfrak{P}$. If we define $\exp y$ to be k and $\mid y \mid = \gamma^{\exp y}$ where $0 < \gamma < 1$ then y is a non-archimedean valuation of $\mathfrak{F}$, i. e.,

$$\mid y_1 y_2 \mid = \mid y_1 \mid \mid y_2 \mid \qquad \mid y_1 + y_2 \mid \leqq \max (\mid y_1 \mid, \mid y_2 \mid).$$

We showed also that $\mathfrak{F}$ is a cyclic algebra over its centrum $\mathfrak{C}$ and the latter is either a $\mathfrak{p}$-adic field [4] or a field of power series in an indeterminate z with coefficients in a finite field. The former case obtains if the characteristic $\chi(\mathfrak{F}) = 0$ and the latter, which we shall call a z-adic field, if $\chi(\mathfrak{F}) = p \neq 0$.

Now if $\mathfrak{C}$ is any commutative $\mathfrak{p}$-adic or z-adic field it may be topologized by means of the neighborhoods $\{a + U_k\}$ of a where U_k is the set of points $(\alpha_0 + \alpha_1\pi + \cdots)\pi^k$, $k = 0, \pm 1, \pm 2, \cdots$, $(\mathfrak{p} = (\pi))$ in the $\mathfrak{p}$-adic field or $(\alpha_0 + \alpha_1 z + \cdots)z^k$ in the z-adic field. In this topology a sequence of elements $a^{(\nu)} = (\alpha_0^{(\nu)} + \alpha_1\pi^{(\nu)} + \cdots)\pi^{k_\nu}$ $((\alpha_0^{(\nu)} + \alpha_1^{(\nu)} z + \cdots)z^{k_\nu})$ converges to 0 if and only if $k_\nu \to \infty$. It follows that the neighborhoods $a + U_k$ are closed and, since the α_i have only a finite range, it is easily seen that they are also compact. The set of elements of finite length, e. g. $(\alpha_0 + \alpha_1\pi + \cdots \alpha_m\pi^m)\pi^k$ is denumerable and everywhere dense in $\mathfrak{C}$. Hence $\mathfrak{C}$ is a l. c. s. t. d. field. If $\mathfrak{F}$ is an algebra with a finite basis over $\mathfrak{C}$ it is a vector space over $\mathfrak{C}$ and hence it is l. c. s. t. d. in the usual vector space topology. Thus we see that the theory of l. c. s. t. d. fields is equivalent to that of division algebras over a $\mathfrak{p}$-adic or a z-adic field. In this note we shall apply the topological methods

* Presented to the Society September 7, 1937. Received by the Editors June 15, 1937.

[1] National Research Fellow.

[2] "Totally disconnected locally compact rings," *American Journal of Mathematics*, vol. 58 (1936), pp. 433-449. This paper will be referred to as T.

[3] An equivalent condition is that $\{a^\nu\}$ have no divergent subsequence: T, p. 444.

[4] i. e. an algebraic extension of finite order of a field of Hensel p-adic numbers.

of our earlier paper to investigate the automorphisms and anti-automorphisms of $\mathfrak{F}$.

Let J be an automorphism or an anti-automorphism of the l. c. s. t. d. field $\mathfrak{F}$, i. e. a $(1-1)$ correspondence of $\mathfrak{F}$ on itself such that $(a+b)^J = a^J + b^J$ and $(ab)^J = a^J b^J$ (automorphism) or $(ab)^J = b^J a^J$ (anti-automorphism). If $c \in$ the centrum $\mathfrak{C}$, $c^J \in \mathfrak{C}$ also and hence the correspondence $c \to c^J$ is an automorphism of $\mathfrak{C}$. By a theorem of F. K. Schmidt [5] J is a continuous transformation in $\mathfrak{C}$.

Suppose $b_1, b_2, \cdots, b_m$ $(m = r^2)$ [6] is a basis for $\mathfrak{F}$ over $\mathfrak{C}$. Then every y may be written uniquely in the form $\Sigma c_i b_i$, $c_i \in \mathfrak{C}$. Set $c_i = E_i(y)$.

LEMMA 1. *$E_i(y)$ is a continuous function of y.*

Since $E_i(y)$ is linear it suffices to show that if $y_\nu = \Sigma c_i^{(\nu)} b_i \to 0$ then $E_i(y_\nu) = c_i^{(\nu)} \to 0$. If $c_i^{(\nu)} \nrightarrow 0$ there exists a subsequence which does not have 0 as a limit point and we may suppose that this is the whole sequence $\{c_i^{(\nu)}\}$. For each ν let $\lambda(\nu) = 1, \cdots, m$ be an index such that $| c_{\lambda(\nu)}^{(\nu)} | \geqq | c_j^{(\nu)} |$ for all $j = 1, \cdots, m$. Since $\lambda(\nu)$ has a finite range, one of its values, say $\lambda(\nu) = 1$, occurs infinitely often and hence by restricting ourselves to a subsequence we may suppose that $| c_1^{(\nu)} | \geqq | c_j^{(\nu)} |$ for all ν. Since $\lim c_i^{(\nu)} \neq 0$, $\lim c_1^{(\nu)} \neq 0$ and there is a subsequence such that $\lim (c_1^{(\nu_k)})^{-1} = d \in \mathfrak{C}$.[7] Again we write ν for ν_k Then

$$\lim (c_1^{(\nu)})^{-1} y_\nu = 0 = \lim (b_1 + (c_1^{(\nu)})^{-1} c_2^{(\nu)} b_2 + \cdots + (c_1^{(\nu)})^{-1} c_m^{(\nu)} b_m).$$

Since $|(c_1^{(\nu)})^{-1} c_j^{(\nu)}| \leqq 1$ we may suppose that $c_1^{(\nu)} c_j^{(\nu)} \to c_j \in \mathfrak{C}$. It follows that $b_1 + c_2 b_2 + \cdots + c_m b_m = 0$ contrary to the linear independence of the b's.

LEMMA 2. *J is a continuous mapping of $\mathfrak{F}$ on itself.*

Suppose

$$y_\nu = c_1^{(\nu)} b_1 + \cdots + c_m^{(\nu)} b_m \to y = c_1 b_1 + \cdots + c_m b_m.$$

By Lemma 1, $c_i^{(\nu)} \to c_i$ and since J is continuous in $\mathfrak{C}$, $c_i^{(\nu)J} \to c_i^J$. Hence

$$y_\nu^J = c_1^{(\nu)J} b_1^J + \cdots + c_m^{(\nu)J} b_m^J \to y^J = c_1^J b_1^J + \cdots + c_m^J b_m^J.$$

---

[5] F. K. Schmidt, "Mehrfach perfekte Körper," *Mathematische Annalen*, vol. 108 (1933), pp. 1-25.

[6] The order of a division algebra over its centrum is the square of an integer called its degree. See, for example, Deuring's "Algebren," *Ergebnisse der Mathematik*, p. 47.

[7] We require here the property of a locally compact field that $a_\nu \to 0$ if $\{a_\nu\}$ diverges: T, p. 442.

Since J^{-1} is also an automorphism or anti-automorphism it is continuous and hence J is a homeomorphism of $\mathfrak{F}$.

THEOREM 1. *Any automorphism or anti-automorphism J of $\mathfrak{F}$ is isometric, i. e., $|y^J| = |y|$ for every y in $\mathfrak{F}$.*

The condition that $\exp u = 0$ is that $\{u^\nu\}$ and $\{u^{-\nu}\}$ be bounded (or have no divergent subsequence). Since J is a homeomorphism $\{(u^J)^\nu\} = \{(u^\nu)^J\}$ and $\{(u^J)^{-\nu}\}$ are bounded also. Hence $\exp u^J = 0$. Now let x be an element such that $\mathfrak{P} = (x)$, or $\exp x = 1$. Consider x^J. Since $(x^J)^\nu = (x^\nu)^J \to 0$, $x^J \in \mathfrak{P}$ and hence $\exp x^J = l > 0$. Similarly $\exp x^{J^{-1}} = k > 0$. If $x^J = ux^l$ where $\exp u = 0$, $x = u^{J^{-1}}(x^{J^{-1}})^l$ or $(x^{J^{-1}})^l u^{J^{-1}}$. Hence $1 = \exp x = lk$ and so $l = k = 1$. Thus $\exp x^J = 1$. If $\exp y = n$ an integer, $y = ux^n$ and $y^J = u^J(x^J)^n$ or $= (x^J)^n u^J$ and hence $\exp y^J = n$. Hence for any y, $|y^J| = |y|$.

The condition that $\mathfrak{F}$ have an automorphism is evidently no restriction since the identity transformation is such a correspondence. On the other hand we shall see that the existence of an anti-automorphism strongly restricts the structure of $\mathfrak{F}$.

We recall that $\mathfrak{F}$ is cyclic over $\mathfrak{C}$. Its degree $r(= \sqrt{m})$ over $\mathfrak{C}$ may be characterized as the order of the automorphism $u \to xux^{-1} \pmod{\mathfrak{P}}$ of the finite field $\mathfrak{J} - \mathfrak{P}$ where x is determined as above.[8] We have for any u in $\mathfrak{J}$

$$xux^{-1} \equiv u^s \pmod{\mathfrak{P}} \qquad (s = p^t)$$

where p is the characteristic of $\mathfrak{J} - \mathfrak{P}$. It follows that

$$\begin{aligned}(x^J)^{-1}u^J x^J &\equiv (u^s)^J \pmod{\mathfrak{P}} \\ &\equiv (u^J)^s \pmod{\mathfrak{P}}.\end{aligned}$$

Since $|x^J| = |x|$, $x^J = vx$ where $|v| = 1$. Then

$$(u^J)^s \equiv x^{-1}(v^{-1}u^J v)x \equiv x^{-1}u^J x \pmod{\mathfrak{P}}$$

since $v^{-1}u^J v \equiv u^J \pmod{\mathfrak{P}}$. Thus

$$x^{-1}u^J x \equiv xu^J x^{-1} \pmod{\mathfrak{P}}$$

for all u^J. Since u^J varies over a complete set of residues mod $\mathfrak{P}$ when u does we have

$$\begin{aligned}x^{-1}ux &\equiv xux^{-1} \pmod{\mathfrak{P}} \\ x^2ux^{-2} &\equiv u \pmod{\mathfrak{P}}.\end{aligned}$$

Thus the order of the automorphism $u \to xux^{-1} \pmod{\mathfrak{P}}$ is either 1 or 2 and hence $r = 1$ or 2.

[8] T, pp. 446-448.

THEOREM 2. *The degree of an anti-automorphic l. c. s. t. d. field is either* 1 *or* 2.

The condition $r = 1$ or 2 is also sufficient that $\mathfrak{F}$ have an anti-automorphism. For if $r = 1$, $\mathfrak{F}$ is commutative and any automorphism is an anti-automorphism. If $r = 2$, $\mathfrak{F}$ is a quaternion algebra over its centrum and it is well known that $\mathfrak{F}$ has anti-automorphisms J. In fact J may be chosen so that $J^2 = I$ the identity and $c^J = c$ for all c in $\mathfrak{C}$.

We suppose now that J is an involutorial anti-automorphism ($J^2 = I$) and the characteristic $\chi(\mathfrak{F}) \neq 2$. Let $\mathfrak{K}_J$ denote the symmetric or invariant elements ($h^J = h$), $\mathfrak{S}_J$ the set of skew elements ($s^J = -s$) and $\mathfrak{C}_0 = \mathfrak{K}_J \cap \mathfrak{C}$. Evidently $\mathfrak{C}_0$ is a closed subfield of $\mathfrak{C}$ and hence is l. c. s. t. d. For any y we have $y = \frac{1}{2}(y + y^J) + \frac{1}{2}(y - y^J) = h + s$ where $h = \frac{1}{2}(y + y^J) \in \mathfrak{K}_J$ and $s = \frac{1}{2}(y - y^J) \in \mathfrak{S}_J$. In particular this holds for y in $\mathfrak{C}$ and in this case $h \in \mathfrak{C}_0$, $s \in \mathfrak{S}_J \cap \mathfrak{C}$. Hence if $\mathfrak{C} \neq \mathfrak{C}_0$ there exists an element $q \neq 0$ in $\mathfrak{S}_J \cap \mathfrak{C}$. Then $y = h + (sq^{-1})q$ and $sq^{-1} \in \mathfrak{K}_J$. For c in $\mathfrak{C}$ we obtain a quadratic extension $\mathfrak{C}_0(q)$ of $\mathfrak{C}_0$. It follows that $\mathfrak{C}_0$ is infinite and therefore is either a $\mathfrak{p}$-adic or a z-adic field.[9] We note also that $q^2 = q_0$ is symmetric

We wish to show that if $\mathfrak{C} \neq \mathfrak{C}_0$ then $\mathfrak{F} = \mathfrak{C}$, i. e., $\mathfrak{F}$ is commutative. For the present we drop the restriction that $\mathfrak{C}_0$ is l. c. s .t. d. and allow it to denote any commutative field of characteristic $\neq 2$. We suppose that $\mathfrak{F}$ is a division algebra of order 8 over $\mathfrak{C}_0$ with centrum $\mathfrak{C} = \mathfrak{C}_0(q)$ and J is an involutorial anti-automorphism of $\mathfrak{F}$ leaving the elements of $\mathfrak{C}_0$ invariant and mapping q into $-q$. The order of $\mathfrak{F}$ over $\mathfrak{C}$ is of course 4.

It is easily seen that $\mathfrak{K}_J$ and $\mathfrak{S}_J$ defined as above are vector spaces over $\mathfrak{C}_0$. If $h \in \mathfrak{K}_J$, $hq \in \mathfrak{S}_J$ and if $s \in \mathfrak{S}_J$, $sq \in \mathfrak{K}_J$. Hence these spaces have the same order over $\mathfrak{C}_0$. Evidently $\mathfrak{S}_J \cap \mathfrak{K}_J = 0$ and as we showed before $\mathfrak{F} = \mathfrak{S}_J + \mathfrak{K}_J$. It follows that $\mathfrak{K}_J$ and $\mathfrak{S}_J$ have order 4 over $\mathfrak{C}_0$ and since $\mathfrak{K}_J = \mathfrak{S}_J q$, $\mathfrak{F} = \mathfrak{S}_J \mathfrak{C}$.[10]

Let $T(y)$ and $N(y)$ denote respectively the trace and norm of y relative to $\mathfrak{C}$. Then $y^2 - T(y)y + N(y) = 0$ and if $y \notin \mathfrak{C}$ this is the equation of least degree having coefficients in $\mathfrak{C}$ and satisfied by y. Then

$$(y^J)^2 - T(y)^J y^J + N(y)^J = 0 = (y^J)^2 - T(y^J)y^J + N(y^J)$$

and hence for $y \notin \mathfrak{C}$

---

[9] We note that any infinite closed subfield $\mathfrak{C}_0$ of a l. c. s. t. d. field is not discrete. For if $\{a_\nu\}$ is a divergent sequence in $\mathfrak{C}_0$ then $a_\nu^{-1} \to 0$. and hence $\in \mathfrak{C}_0$.

[10] If $\mathfrak{A}$ and $\mathfrak{B}$ are vector spaces over $\mathfrak{C}_0$, $\mathfrak{AB}$ denotes the smallest vector space over $\mathfrak{C}_0$ containing all the products ab where $a \in \mathfrak{A}$ and $b \in \mathfrak{B}$.

$$(*) \qquad T(y^J) = T(y)^J \qquad N(y^J) = N(y)^J.$$

But for $y \in \mathfrak{C}$, $T(y) = 2y$, $N(y) = y^2$ and so (*) holds for every y in $\mathfrak{F}$. If $y \in \mathfrak{S}_J$, $T(y)^J = T(y^J) = T(-y) = -T(y)$ and $N(y)^J = N(-y) = N(y)$, i. e., $T(y) \in \mathfrak{S}_J \frown \mathfrak{C}$ and $N(y) \in \mathfrak{C}_0$. $\mathfrak{S}_J \frown \mathfrak{C}_0$ consists of the $\mathfrak{C}_0$-multiples of q.

Let $\mathfrak{S}'_J$ be the totality of elements of trace 0 in $\mathfrak{S}_J$. $\mathfrak{S}'_J$ is a vector space over $\mathfrak{C}_0$ and if y, $t \in \mathfrak{S}'_J$ so does $[y, t] = yt - ty$. If y is any element of $\mathfrak{S}_J$, $y = \frac{1}{2}T(y) + y_0$ where $y_0 = y - \frac{1}{2}T(y) \in \mathfrak{S}'_J$. Thus $\mathfrak{S}_J = \mathfrak{S}'_J + \mathfrak{C}_0 q$ and $\mathfrak{S}'_J \frown \mathfrak{C}_0 q = 0$. It follows that the order of $\mathfrak{S}'_J$ over $\mathfrak{C}_0$ is 3. Since $1 \nmid \mathfrak{S}'_J$, the order of $\mathfrak{F}_0 = \mathfrak{S}'_J + \mathfrak{C}_0 1$ is 4 over $\mathfrak{C}_0$.

For y, t in $\mathfrak{S}'_J$ we have

$$\begin{aligned} yt &= \tfrac{1}{2}[(y+t)^2 - y^2 - t^2 + yt - ty] \\ &= \tfrac{1}{2}[-N(y+t) + N(y) + N(t) + yt - ty] \in \mathfrak{F}_0. \end{aligned}$$

Hence $\mathfrak{F}_0$ is an algebra over $\mathfrak{C}_0$. Since $\mathfrak{F}_0 = \mathfrak{S}'_J + \mathfrak{C}_0 1$, $\mathfrak{S}_J = \mathfrak{S}'_J + \mathfrak{C}_0 q$ and $\mathfrak{F} = \mathfrak{S}_J \mathfrak{C}$, we have $\mathfrak{F} = \mathfrak{F}_0 \mathfrak{C}$. We have therefore proved the following theorem.

THEOREM 3. *If $\mathfrak{F}$ is a division algebra of order 8 over $\mathfrak{C}_0$ of characteristic $\neq 2$ with $\mathfrak{C} = \mathfrak{C}_0(q)$, $q^2 \in \mathfrak{C}_0$, as centrum and J an involutorial anti-automorphism in $\mathfrak{F}$ such that $q^J = -q$ then $\mathfrak{F}$ is a direct product $\mathfrak{C} \times \mathfrak{F}_0$ of $\mathfrak{C}$ and $\mathfrak{F}_0$ the join of $\mathfrak{C}_0 1$ and the space of skew elements of trace 0.*[11]

Since the centrum of $\mathfrak{F}$ is $\mathfrak{C}$ that of $\mathfrak{F}_0$ must be $\mathfrak{C}_0$.

If $\mathfrak{C}_0$ is a commutative l. c. s. t. d. field it is well known that an algebra $\mathfrak{F}_0$ of degree 2 over $\mathfrak{C}_0$ has every quadratic field as a splitting field, i. e., if $\mathfrak{C} = \mathfrak{C}_0(q)$ then $\mathfrak{C} \times \mathfrak{F}_0$ is a complete matrix algebra over $\mathfrak{C}$.[12] Hence

THEOREM 4. *If $\mathfrak{F}$ is an l. c. s. t. d. field of characteristic $\neq 2$ and J an involutorial anti-automorphism such that $c^J \neq c$ for some c in the centrum $\mathfrak{C}$, then $\mathfrak{F} = \mathfrak{C}$ is commutative.*

In the remainder of the paper we shall extend the above results to $\mathfrak{F}_n$ a complete matrix algebra of n rows and columns over an l. c. s. t. d. field $\mathfrak{F}$. Let e_{ij} $(i, j = 1, \cdots, n)$ be a set of matrix units in $\mathfrak{F}_n$; $e_{ij}e_{kl} = \delta_{jk}e_{il}$ and every element b in $\mathfrak{F}_n$ has the form $\Sigma y_{ij}e_{ij}$ where $y_{ij} \in \mathfrak{F}$. As before let $\mathfrak{C}$

[11] Cf. A. A. Albert, "Involutorial simple algebras and real Riemann matrices," *Annals of Mathematics*, vol. 36 (1935), p. 909.

[12] Deuring's *Algebren*, p. 113. The proof given there for $\mathfrak{C}_0$ a $\mathfrak{p}$-adic field holds also for a z-adic field. See *Algebren*, p. 137.

denote the centrum of $\mathfrak{F}$ or of $\mathfrak{F}_n$. We recall that $\mathfrak{F}$ may be characterized as the totality of elements commutative with the e_{ij}.

Let J be an anti-automorphism of $\mathfrak{F}_n$ and $e_{ji}{}^J = f_{ij}$. Then $f_{ij}f_{kl} = \delta_{jk}f_{il}$ and it follows that there exists a non-singular element s in $\mathfrak{F}_n$ such that $e_{ij} = s^{-1}f_{ij}s$. The correspondence $b \to b^K = s^{-1}b^Js$ is an anti-automorphism also and $e_{ij}{}^K = e_{ji}$. Hence if $y \in \mathfrak{F}$, y^K which commutes with the e_{ij} belongs to $\mathfrak{F}$ also and $\mathfrak{F}$ itself is anti-automorphic. By Theorem 1 we have

THEOREM 5. *If $\mathfrak{F}$ is an l. c. s. t. d. field such that $\mathfrak{F}_n$ is anti-automorphic then $\mathfrak{F}$ has degree 1 or 2 over its centrum.*

If J is involutorial, then

$$e_{ij} = e_{ij}{}^{J^2} = f_{ji}{}^J = (se_{ji}s^{-1})^J = (s^{-1})^Jse_{ij}s^{-1}s^J.$$

Thus $s^{-1}s^J$ commutes with all e_{ij} and $s^J = su$ where $u \in \mathfrak{F}$. If $u = -1$ it is easily seen that the correspondence $b \to b^K = s^{-1}b^Js$ is an involutorial anti-automorphism in $\mathfrak{F}_n$ mapping $\mathfrak{F}$ into itself. If $u \neq 1$, $s^J + s = s(u+1) = t$ is non-singular and $t^J = t$. Hence the correspondence $b \to b^K = t^{-1}b^Jt$ is an involutorial anti-automorphism mapping $\mathfrak{F}$ into itself.[13] In either case $\mathfrak{F}$ is involutorial anti-automorphic. Evidently if $c^J \neq c$ then $c^K \neq c$ for c in $\mathfrak{C}$. Hence by Theorem 4 we have

THEOREM 6. *If $\mathfrak{F}$ is an l. c. s. t. d. field of characteristic $\neq 2$ and $\mathfrak{F}_n$ has an involutorial anti-automophism J mapping an element c of the centrum into $c^J \neq c$ then $\mathfrak{F} = \mathfrak{C}$ is commutative.*

PRINCETON UNIVERSITY AND
THE INSTITUTE FOR ADVANCED STUDY.

[13] Cf. Albert, *loc. cit.* in 11, p. 897.

THE QUATERNIONIC CONGRUENCE $\bar{\imath}at \equiv b \pmod{g}$ AND THE EQUATION $h(8n+1) = x_1^2 + x_2^2 + x_3^2$.*

By GORDON PALL.

1. A quaternion will be called integral if its coördinates are rational integers. Such quaternions will be represented by the letters $a, \cdots, e$, $t, \cdots, z$, and their integer coördinates will be distinguished by subscripts; thus $t = t_0 + it_1 + jt_2 + kt_3$.

If g is a rational integer, the congruence $c \equiv d \pmod{g}$ is equivalent to the system $c_f \equiv d_f \pmod{g}$ $(f = 0, 1, 2, 3)$. Most of the familiar rules for handling congruences can be extended, with due regard to lack of commutativity, to such congruences. For example, if $a \equiv b$ and $c \equiv d \pmod{g}$, then $ac \equiv bd$; for $ac - bd = (a-b)c + b(c-d)$. Again, if $Nx (= \bar{x}x = \Sigma x_f^2)$ is prime to g, $ax \equiv bx$ implies $a \equiv b$; for $(ax - bx)\bar{x} = (a-b)Nx$.

Some of our results (e. g. § 2 and (19)) will bear incidentally on cancellation of an inner factor. For instance, by (12) either of $\bar{\imath}it \equiv \pm i \pmod{4}$ implies $\bar{\imath}t \equiv 1 \pmod{4}$.

The study of the congruence $\bar{\imath}at \equiv b \pmod{g}$ is reducible to the case where a and b are pure quaternions (§ 5). If $g = 2^n$ and $4 \nmid Nb$, the conditions for mere solvability are simple (Theorems 2 and 4). But all solutions t satisfy one and only one of $\bar{\imath}t \equiv 1$ or $3 \pmod{4}$, as set forth in Theorems 3 and 5.

Our applications to the equation

$$h(8n+1) = x_1^2 + x_2^2 + x_3^2 \tag{1}$$

employ properties of the residues $\bar{\imath}at$ for moduli g dividing $4Na$, where a is any given pure quaternion $ia_1 + ja_2 + ka_3$. Analogous results, in part, exist for any g but will not be treated here.

For any g and (pure) a we obtain a system of residues $\bar{\imath}at \pmod{g}$ as t ranges over all integral quaternions. We regard the quaternions formed from a by permuting or changing the signs of a_1, a_2, a_3 as constituting a *set* $\Sigma(a)$. If a' is so formed by an even number of permutations and an even number of sign-changes, then for any odd modulus g, the residue-systems $\bar{\imath}at$ and $\bar{u}a'u$ are identical (§ 11, end). Thus there are associated with $\Sigma(a)$ at most four residue-systems $(\mathrm{mod}\ g)$, g being odd, namely those represented by

* Received June 18, 1937.

$$\bar{\imath}at,\ \bar{\imath}(-a)t,\ \bar{\imath}a^*t,\ \bar{\imath}(-a^*)t, \tag{1'}$$

where $a^* = ia_1 + ja_3 + ka_2$. The four need not be distinct (Th. 9).

If $\bar{\imath}at \equiv b \pmod g$, $Nb \equiv Na(Nt)^2 \pmod g$. We now assume $g \mid Na$, g odd, a proper and pure. If g is a prime p, then $\bar{\imath}at$ represents besides $0 \pmod p$ exactly half the remaining $p^2 - 1$ residues b such that $p \mid Nb$; if $p \not\equiv 1 \pmod 8$ the other half are represented by $\bar{\imath}(-a)t$, $\bar{\imath}a^*t$, or $\bar{\imath}(-a^*)t$ (§ 14). These results are extended to composite g (§ 14). If g contains a non-square factor $\equiv 1 \pmod 8$ not all residues b of norm divisible by g are represented by (1'). It may be possible to supply the deficiency by other quaternions of the same norm as a, but this we shall not here investigate.

In (1) h and n are assumed to be non-negative integers, $4 \nmid h$. If $8n + 1$ is a square m^2, $m > 0$, then all proper integral solutions $x = ix_1 + jx_2 + kx_3$ of (1) are of the form $\bar{\imath}at$, where $a = ia_1 + ja_2 + ka_3$ is an integral solution of

$$h = a_1^2 + a_2^2 + a_3^2 \tag{2}$$

and t is a proper quaternion of norm m (*proper* meaning having coprime coördinates). To prove this we use the following well-known and fundamental result in the arithmetic of quaternions:

A proper quaternion, with norm divisible by an odd positive integer m, has exactly eight right-divisors (left-divisors) t of norm m. These form a class of left-associate (right-associate) quaternions $\pm t$, $\pm it$, $\pm jt$, $\pm kt$ ($\pm \bar{\imath}$, $\pm \bar{\imath}i$, $\pm \bar{\imath}j$, $\pm \bar{\imath}k$).

Consider then a proper, pure quaternion x of norm hm^2. Its left-divisors of norm m (unique up to a right unit factor) are the conjugates of the right-divisors. For if $x = vt$, $\bar{x} = -x = \bar{\imath}\bar{v}$. Hence, since $Nv = hm$, $\bar{\imath}$ is a left-divisor of v, $v = \bar{\imath}a$, $x = \bar{\imath}at$.

If x is derived in this way from a, $\mathfrak{S}(x)$ is derived from $\mathfrak{S}(a)$. For any proper, pure x of norm hm^2, $x = \bar{\imath}at$, where $Nt = m$, and t is unique up to a left unit factor. If t is replaced by $\pm it$, a is replaced by $-iai = ia_1 - ja_2 - ka_3$; similarly for $\pm jt$ and $\pm kt$. Evidently $-x$ is derived from $-a$; and if x_1, x_2, x_3 be permuted or changed in sign, so that $x \to x'$, then x' is derived from a' similarly obtained (cf. § 9, near end). Thus $\mathfrak{S}(x)$ is derived from $\mathfrak{S}(a)$. In § 15 we shall complete the proof of

THEOREM 1. *Let $h > 0$, $4 \nmid h$. Let h' be a positive divisor of h such that $h/h' \equiv 1 \pmod 8$ (e. g. $h' = h$). Let A denote the class of all residues $\pmod{4h'}$ of proper solutions x of (1) derived from a single proper set of solutions of (2), with $8n + 1 = m^2$ ($m = 1, 5, 9, 13, \cdots$); and let B denote*

the class of residues (mod $4h'$) *obtained similarly with* $m = 3, 7, 11, 15, \cdots$. *Then:*

I. the classes A and B are mutually exclusive; and

II. if h' has no factor > 1 of the form $8f + 1$, then all proper solutions of (1) *belong to A or B, and for any n such that $(8n + 1)h/h'$ is not a square,* (1) *has equally many solutions in each class A, B.*

Whenever (2) has only one proper set of solutions, e. g. when $h < 17$, the clause "derived from . . . of (2)" can be omitted. The classes A and B are then easily calculated by forming solutions of $h(4f \pm 1)^2 = x_1^2 + x_2^2 + x_3^2$, the work being abbreviated by using § 16. We give a table, listing only the least absolute residues (mod $4h$) of (x_1, x_2, x_3), and arranging each triple in a definite order:

h	residues A, $m = 4f + 1$	residues B, $m = 4f + 3$
1	(1, 0, 0)	(1, 2, 2)
2	(1, 1, 0), (3, 3, 4)	(3, 3, 0), (1, 1, 4)
3	(1, 1, 1), (1, 5, 5)	(1, 1, 5), (5, 5, 5)
5	(1 or 9, 0, 2), (3 or 7, 4, 10), (5, 8, 6)	(9 or 1, 10, 8), (7 or 3, 6, 0), (5, 2, 4)
6	(2, 1, 1), (2, 7, 7), (2, 5, 11)	(10, 1, 1), (10, 7, 7), (10, 5, 11)
	(10, 5, 5), (10, 11, 11), (10, 1, 7)	(2, 5, 5), (2, 11, 11), (2, 1, 7)
10	(0, 1 or 9, 3 or 13), (20, 7 or 17, 11 or 19)	change the even
	(4, 3 or 13, 15), (16, 7 or 17, 5)	or odd x_i
	(8, 11 or 19, 15), (12, 1 or 9, 5)	of A
11	(1, 1, 3), (3, 3, 13), (5, 5, 15), (7, 7, 21), (9, 9, 17)	to
	(21, 21, 3), (19, 19, 13), (17, 17, 15), (15, 15, 21), (13, 13, 17)	$2h - x_i$
	(1, 21, 19), (3, 19, 9), (5, 17, 7), (7, 15, 1), (9, 13, 5)	according as
13	(0, 2, 3 or 23), (20, 2, 5 or 21), (4, 14, 3 or 23),	$h \not\equiv 3$ or
	(12, 26, 5 or 21), (8, 14, 13), and multiply	$h \equiv 3 \pmod 8$.
	these by 3 and 9, reducing mod 52.	

For example let $h = 2$. If $8n + 1$ is a square $(4f + 1)^2$, $f \geqq 0$, the proper solutions of (1) have the two odd $x_i \equiv \pm 1 \pmod 8$ if the even x_j is $\equiv 0 \pmod 8$, but the two odd x_i are $\equiv \pm 3 \pmod 8$ if the even x_j is $\equiv 4 \pmod 8$; vice versa if $8n + 1 = (4f + 3)^2$. But if $8n + 1$ is not a square then $2(8n + 1)$ has equally many proper representations of either type. Thus $34 = 5^2 + 3^2 + 0^2 = 3^2 + 3^2 + 4^2$.

The case $h = 2$ can be given a more elegant form:

In the equation $2(8n + 1) = x_1^2 + x_2^2 + x_3^2$ (x_3 even), there are equally many solutions of each type

$$P: \quad x_1 \equiv \pm x_2 \pmod{16}, \qquad Q: \quad x_1 \equiv \pm x_2 + 8 \pmod{16},$$

when $8n+1$ is not a square. If $8n+1 = s^2, s > 0$, then all proper solutions are of type P if $s \equiv 1$ or $3 \pmod 8$, but of type Q if $s \equiv 5$ or $7 \pmod 8$.

For A yields P if n is even, Q if n is odd; vice versa for B.

Hence $8n+1 = y_1^2 + y_2^2 + 2y_3^2$, where

$$y_1 = \tfrac{1}{2}(x_1 + x_2), \; y_2 = \tfrac{1}{2}(x_1 - x_2), \; x_3 = 2y_3.$$

As a corollary of cases $h = 1$ and 2, every $8n+1$ is represented in both forms

$$x_1^2 + 16x_2^2 + 16x_3^2, \quad x_1^2 + 8x_2^2 + 64x_3^2;$$

whence both forms are easily shown to be regular, although their genera contain two classes. The present theory has extensions to many ternary quadratic forms.

As a further corollary, every $8n+1$ except 1 and 25, and every $2(8n+1)$ except 2, is a sum of three positive squares.[1]

For $h = 1$ and 3 Theorem 1 is essentially equivalent to certain results of Jacobi and Glaisher obtained by theta-function expansions. For references, see Dickson's *History*, II, to Jacobi, pp. 262-3; Catalan, p. 266; Glaisher, p. 268; Bachmann, p. 270. Glaisher conjectured (*Quarterly Journal of Mathematics*, 20 (1885), p. 96) that his results could be proved by "actual transformation of the squares," and they were in fact proved recently by the writer by using automorphs of $x_1^2 + x_2^2 + x_3^2$, but not published. The close connection between such automorphs and quaternions is well known.

2. Even if Na is prime to g, $\bar{\imath}at \equiv a \pmod g$ need not imply $\bar{\imath}t \equiv 1 \pmod g$. For example, if g is *odd* and $a = i$, $\bar{\imath}it \equiv i$ expands into

$$t_0^2 + t_1^2 - t_2^2 - t_3^2 \equiv 1, \quad -t_0t_3 + t_1t_2 \equiv 0, \quad t_0t_2 + t_1t_3 \equiv 0.$$

These are compatible with $\bar{\imath}t \equiv -1$ if $0 \equiv t_0 \equiv t_1 \equiv 1 + t_2^2 + t_3^2$, and with $\bar{\imath}t \equiv 1$ if $0 \equiv t_2 \equiv t_3 \equiv t_0^2 + t_1^2 - 1$, either set of conditions being solvable.

That $\bar{\imath}t \equiv \pm 1$ are the only possibilities is shown in § 3, provided g is the power of a prime. We assume that g is the power of a prime throughout § 3.

3. On taking norms. From $\bar{\imath}at \equiv b \pmod g$ follows $\bar{\imath}at = b + gv$, $\bar{\imath}\bar{a}\bar{t} = \bar{b} + g\bar{v}$, $\bar{\imath}at\bar{\imath}\bar{a}\bar{t} = b\bar{b} + g(b\bar{v} + v\bar{b}) + g^2v\bar{v}$, i. e.

[1] This was conjectured by the writer in "On Sums of Squares," *American Mathematical Monthly*, XL (1933), 10-18, on p. 11. The number 137 should be omitted from Theorem 3 of that article.

$$(Nt)^2 Na - Nb = g(b\bar{v} + v\bar{b}) + g^2 Nv. \tag{3}$$

Here $b\bar{v} + v\bar{b}$, a quaternion plus its conjugate, is an even integer.

Assuming $Na \equiv Nb \pmod{2g}$ we draw some conclusions from (3). If Na is prime to g and g is odd, $(Nt)^2 \equiv 1$ and $Nt \equiv \pm 1 \pmod{g}$. We have just observed that both residues ± 1 are obtainable for Nt when $a = i$ and g is odd. Although, as we shall see for example in (19), the like result does not hold for even g, (3) yields only $Nt \equiv \pm 1 \pmod{g}$ if Na is prime to g (g even); and $Nt \equiv \pm 1 \pmod{\frac{1}{2}g}$ if $4 \mid g$, $Na \equiv 2 \pmod 4$, $\frac{1}{2}Na$ prime to g.

Conversely, if neither Na nor Nb is divisible by 4, and $2 \mid g$, (3) requires Nt odd, $Na \equiv Nb \pmod 4$ if $2 \mid g$, $Na \equiv Nb \pmod 8$ if $4 \mid g$. And if $8 \mid g$, and $Na \equiv Nb \equiv \pm 2 \pmod 8$, (3) implies $Na \equiv Nb \pmod{16}$.

4. On expanding $\bar{\imath}at$ (cf. (42) and § 5) we see that

$$\bar{\imath}at \equiv (Nt)a \pmod 2, \tag{4}$$

for any integral quaternions t, a. Hence if $2 \nmid b$,

$$\bar{\imath}at \equiv b \pmod 2 \text{ implies } a \equiv b \pmod 2.$$

5. Henceforth we shall treat $\bar{\imath}at \equiv b \pmod g$ with $a = ia_1 + ja_2 + ka_3$, $b = ib_1 + jb_2 + kb_3$ (pure quaternions). Nothing essential is hereby lost. For if $x = a_0 + a$, $y = b_0 + b$, then the real part of $\bar{\imath}xt$ is $\bar{\imath}a_0 t = a_0 Nt$, and $\bar{\imath}xt \equiv y$ breaks up into

$$a_0 Nt \equiv b_0, \qquad \bar{\imath}at \equiv b.$$

6. Principal theorems for $g = 2^n$. Hypothesis throughout:

$$a = ia_1 + ja_2 + ka_3,\ b = ib_1 + jb_2 + kb_3 \tag{5}$$

are pure integral quaternions, neither divisible by 2;

$$l = 1 \text{ if } Na \text{ is odd}, \qquad l = 2 \text{ if } Na \equiv 2 \pmod 4. \tag{6}$$

Like a familiar theorem on quadratic congruences in elementary theory of numbers is the following:

THEOREM 2. *The solvability for t, necessarily of odd norm, of*

$$\bar{\imath}at \equiv b \pmod{4l} \tag{7}$$

implies, for every integer $n > 0$, the solvability for t of

$$\bar{\imath}at \equiv b \pmod{2^n}. \tag{8}$$

However a new feature is presented in

THEOREM 3. *If* (7) *is solvable, then all solutions* t *satisfy one and only one of the congruences*

$$(9) \qquad A: \ \bar{\imath}t \equiv 1 \pmod 4, \qquad B: \ \bar{\imath}t \equiv -1 \pmod 4.$$

THEOREM 4. *Congruence* (7) *is solvable for* t *if and only if*

$$(10) \qquad a \equiv b \pmod 2, \qquad Na \equiv Nb \pmod{8l}.$$

THEOREM 5. *Conditions* (10) *being assumed, the solutions of* (7) *are of type* A *or* B *in accordance with the following rules:*

(11) *if* $Na \equiv 3 \pmod 8$, A *holds or* B *according as* $a_h \equiv -b_h \pmod 4$ *for an even* (0 *or* 2) *or odd* (1 *or* 3) *number of values* $h = 1, 2, 3$;

(12) *if* $Na \equiv 1 \pmod 4$, *and* a_1 (*say*) *is odd, then* A *holds if* $a_2 \equiv b_2$, $a_3 \equiv b_3$, B *if* $a_2 \equiv b_2 + 2$, $a_3 \equiv b_3 + 2 \pmod 4$;

(13) *if* $Na \equiv 2 \pmod 8$ *and* $4 \mid a_3$, A *holds if* $(b_1, b_3) \equiv (\pm a_1, a_3)$ *or* $(\pm 3a_1, a_3 + 4) \pmod 8$, B *if* $(b_1, b_3) \equiv (\pm a_1, a_3 + 4)$ *or* $(\pm 3a_1, a_3)$;

(13′) *if* $Na \equiv 6 \pmod 8$ *and* $a_3 \equiv b_3 \equiv \pm 2$, A *holds if* $\pm (b_1, b_2) \equiv (a_1, a_2)$ *or* $3(a_1, -a_2)$, B *if* $\pm (b_1, b_2) \equiv (a_1, -a_2)$ *or* $3(a_1, a_2)$; A *and* B *being reversed if* $a_3 \equiv -b_3 \pmod 8$.

The subscripts 1, 2, 3 *may be permuted in* (12)-(13′).

Theorem 2 is proved in § 7, 3 in § 8, 4 and 5 in §§ 9, 10.

7. Proof of Theorem 2. Proceeding by induction we may suppose that $\bar{\imath}at - b = 2^n w$, where $w = iw_1 + jw_2 + kw_3$ is a pure, integral quaternion, $n > l$.

First take Na to be odd, $n \geqq 2$, and set

$$(14) \qquad u = t + 2^{n-1}z, \ z \text{ an integral quaternion.}$$

Then $\bar{u}au - b = 2^n e$, where

$$e = w + \tfrac{1}{2}(\bar{z}at + \bar{\imath}az) + 2^{n-2}\bar{z}az,$$

and we wish to choose z to make e even. As in (4), $\bar{z}az \equiv (Nz)a \pmod 2$, so that it suffices to have $\frac{1}{2}(\bar{\imath}az + \bar{z}at) \equiv w \pmod 2$ and Nz even. Abbreviate $\bar{\imath}a = d = d_0 + id_1 + \cdots$, whence $at = -\bar{d}$ and Nd is odd. The desired conditions expand into

$$
\begin{aligned}
&d_1z_0 + d_0z_1 - d_3z_2 + d_2z_3 \equiv w_1,\\
&d_2z_0 + d_3z_1 + d_0z_2 - d_1z_3 \equiv w_2,\\
&d_3z_0 - d_2z_1 + d_1z_2 + d_0z_3 \equiv w_3,\\
&-z_0 + z_1 - z_2 + z_3 \equiv 0, \pmod 2.
\end{aligned} \tag{15}
$$

The determinant of the coefficients in the left members is

$$Nd(d_0 + d_1 + d_2 + d_3).$$

This being odd, conditions (15) can be satisfied.

Second take $Na \equiv 2 \pmod 4$, $n \geqq 3$, and set

$$u = t + 2^{n-2}\theta z, \tag{16}$$

where $\theta = 1 + i$, $1 + j$, or $1 + k$ according as a_1, a_2, or a_3 is even. Then $\bar{\theta}a = 2v$, where v is an integral quaternion of odd norm. Now $\bar{u}au - b = 2^n e$, where $e = w + \frac{1}{2}(\bar{z}vt - \bar{t}\bar{v}z) + 2^{n-3}\bar{z}v\theta z$. Since Nv is odd, the argument proceeds as in the first case.

8. We prove a more general result than Theorem 3:

THEOREM 6. *Assume* (5) *and* (6), $n \geqq 2$. *All solutions* t *of*

$$\bar{\imath}at \equiv b \pmod{l \cdot 2^n}, \tag{17}$$

if any exist, satisfy one and only one of

$$\bar{\imath}t \equiv 1 \pmod{2^n}, \qquad \bar{\imath}t \equiv -1 \pmod{2^n}, \tag{18}$$

the particular one depending only on the residues of a *and* $b \pmod{4l}$.

It is expedient first to prove:

$$\bar{\imath}at \equiv a \pmod{l \cdot 2^n} \text{ implies } \bar{\imath}t \equiv 1 \pmod{2^n}; \tag{19}$$

(5) and (6) being assumed. Since $\bar{\imath}at \equiv a$, $(Nt)at \equiv ta$, i. e.

$$Nt(a_1t_1 + a_2t_2 + a_3t_3) \equiv a_1t_1 + a_2t_2 + a_3t_3, \tag{20}$$

$$Nt(a_1t_0 + a_2t_3 - a_3t_2) \equiv a_1t_0 - a_2t_3 + a_3t_2, \tag{21}$$

$$Nt(-a_1t_3 + a_2t_0 + a_3t_1) \equiv a_1t_3 + a_2t_0 - a_3t_1, \tag{22}$$

$$Nt(a_1t_2 - a_2t_1 + a_3t_0) \equiv -a_1t_2 + a_2t_1 + a_3t_0, \tag{23}$$

to modulus $l \cdot 2^n$. Form $(21)a_1 + (22)a_2 + (23)a_3$:

$$(Nt)t_0(a_1^2 + a_2^2 + a_3^2) \equiv t_0(a_1^2 + a_2^2 + a_3^2), \pmod{l \cdot 2^n}. \tag{24}$$

The conclusion $\bar{t}t \equiv 1 \pmod{2^n}$ follows from (20) or (24), if

(25) $\qquad a_1t_1 + a_2t_2 + a_3t_3$ or t_0 is odd.

In fact (25) always holds in consequence of (20)-(23): to see this, we prove the impossibility of

(α) t_0 even, t_1, t_2, t_3 odd, a_1, a_2 odd, a_3 even;
(β) t_0, t_1, t_2 even, t_3 odd, a_3 even.

In case (α), $Na \equiv 2 \pmod 4$, and (21) holds to modulus 8. If $4|t_0$, $Nt \equiv 3 \pmod 8$, (21) reduces to $4 \equiv 0 \pmod 8$. If $4|t_0 - 2$, $Nt \equiv -1 \pmod 8$, (21) implies $8|2a_1t_0$. In case (β), $Nt \equiv 1 \pmod 4$, and (21) yields $4|2(a_2t_3 - a_3t_2)$, a contradiction unless a_2 is even. But then a_1 is odd, and (22) yields $4|2(a_3t_1 - a_1t_3)$, whereas $a_3t_1 - a_1t_3$ is odd.

Now Theorem 6 will follow from Theorem 3. For by § 3, (17) implies $\bar{t}t \equiv \pm 1 \pmod{2^n}$, and only one of these agrees with either residue $\pm 1 \pmod 4$.

If, besides (7) we have $\bar{u}au \equiv b \pmod{4l}$, then $(Nu)^2a \equiv ub\bar{u}$, $u\bar{t}at\bar{u} \equiv a \pmod{4l}$, and by (19) with $n = 2$, $N(t\bar{u}) = NtNu \equiv 1 \pmod 4$, $Nt \equiv Nu \pmod 4$.

9. Let us write aAb to indicate that (7) is solvable with $\bar{t}t \equiv 1$, and aBb to express solvability with $\bar{t}t \equiv -1 \pmod 4$. Let S denote either A or B, but a fixed one of them on a given line. Trivially,

(26) $\qquad aSb$ implies $a'Sb'$,

if $a' \equiv ha$, $b' \equiv hb$, $\pmod{4l}$, h an odd integer;

(27) $\qquad aSb$ implies bSa;

(28) $\qquad aAa$;

(29) $\qquad aSb$, bSc, imply aAc;

(30) $\qquad aAb$, bBc, imply aBc.

If $a'' = ia_1 - ja_2 - ka_3$, $-ia_1 + ja_2 - ka_3$, or $-ia_1 - ja_2 + ka_3$ is obtained from a by

(31) $\qquad$ changing the signs of any two of the a_i,

then aAa'', and

(31′) $\qquad aSb$ implies $a''Sb$, aSb'', $a''Sb''$.

For $a'' = -iai$, $-jaj$, or $-kak$, and $N(ti) = Nt$, etc.

The third conclusion in (31′) may be generalized: *if the same sequence of permutations and sign-changes be applied to the coördinates of both a and b to produce $a^\dagger$ and $b^\dagger$, then*

(32) $$aSb \text{ implies } a^\dagger S b^\dagger.$$

This is evident as regards an even number of permutations and sign-changes: for an even number of such operations on i, j, k produces a permutation of the entire class of integral quaternions which preserves the truth of all relations constructed by addition, multiplication, etc. If an odd number of operations are involved in forming $a^\dagger$ we note that aSb implies $(-a)S(-b)$, by (26), and use $-a^\dagger$ and $-b^\dagger$ in (32).

Examples: Write (x_1, x_2, x_3) for $ix_1 + jx_2 + kx_3$. From $(1,1,1)A(1,1,1)$ follows $(1,1,1)A(1,-1,-1)$, etc. From $(1,1,2)B(1,1,-2)$ follows by (26) $(3,3,-2)B(3,3,2)$, $(3,-3,-2)B(3,-3,2)$, etc.

10. The necessity of (10) was observed in §§ 3 and 4.

Both (10) and § 9, especially (32), reduce considerably the number of pairs of residue triples to be considered, first as to whether they can be transformed into each other, (mod 4) if of odd norm, (mod 8) if of norm $4n+2$; second, whether the transformation is of type A or B. All cases can be reduced to (33)-(38) below.

To begin with, suppose a_1, a_2, a_3 odd. We have

(33) $$(1,1,1)A(1,1,1), \qquad (1,1,1)B(-1,-1,-1).$$

For (33_2) follows from

$$(-i-j-k)(i+j+k)(i+j+k) \equiv -i-j-k \pmod{4}.$$

Changing any two signs at once we have the parts of Theorems 4 and 5 which concern $Na \equiv 3 \pmod{8}$.

Next let a_1 be odd, a_2 and a_3 even. We have

(34) $$(1,0,0)B(1,2,2), \qquad (1,2,0)B(1,0,2).$$

For

$$(1-i-j)i(1+i+j) = i+2j+2k,$$
$$(1-i-j)(i+2j)(1+i+j) = 5i+4j-2k \equiv i+2k \pmod{4}.$$

The parts of Theorems 4, 5 concerning $Na \equiv 1 \pmod{4}$ now follow.

We prove next the sequences of relations

(35) $$(1,1,0)A(3,3,4)B(1,1,4)A(3,3,0),$$

(36) $$(1,3,0)A(3,1,4)B(1,3,4)A(3,1,0):$$
$$(2i-k)(i+j)(2i+k) = 3i-5j+4k \equiv 3i+3j+4k \pmod 8,$$
$$(2-i-j-k)(i+j)(2+i+j+k) = 9i+j+4k \equiv i+j+4k,$$
$$(1-i-j)(i+j)(1+i+j) = 3i+3j;$$
$$(-2i-k)(i+3j)(2i+k) = 3i-15j+4k \equiv 3i+j+4k,$$
$$(i-j+k)(i+3j)(-i+j-k) = -7i-5j-4k \equiv i+3j+4k.$$

Hence, for example, $(1,1,0)B(1,1,4)$ by (30); none of the residue triples in (35) can be transformed into any in (36), in view of (10). Similarly

(37) $$(1,1,2)A(3,3,-2)B(1,1,-2)A(3,3,2),$$

(38) $$(1,3,2)A(3,1,-2)B(1,3,-2)A(3,1,2):$$
$$(-2i-k)(i+j+2k)(2i+k) = 11i-5j-2k \equiv 3i+3j-2k,$$
$$(-i+j-k)(i+j+2k)(i-j+k) = i-7j-2k \equiv i+j-2k;$$
$$(-2i-k)(i+3j+2k)(2i+k) = 11i-15j-2k \equiv 3i+j-2k,$$
$$(1+j+k)(i+3j+2k)(1-j-k) = (-3i+j-2k)A(3i+j+2k).$$

From (35)-(38), Theorems 4 and 5 may now be easily verified for $l=2$.

11. Part I of Theorem 1 will follow from the case $\lambda = 1$ of

THEOREM 7. *Let a satisfy (2), (5), and (6), and let $a^\dagger$ be obtained from a by merely permuting or changing signs of a_1, a_2, a_3. Let λ be an odd integer prime to $h = Na$. Let H be any positive odd factor of h such that $h/(lH) \equiv 1 \pmod 8$, e. g. $H = h/l$. Then if the g. c. d. of the components of $\bar{\imath}at$ is prime to H, and t, u are of odd norms,*

(39) $$\bar{\imath}at \equiv \lambda \bar{u}a^\dagger u \pmod{4lH} \text{ implies } Nt \equiv (lH|\lambda)Nu \pmod 4.$$

We first extract from § 10 what pertains to $b \equiv \lambda a^\dagger \pmod{4l}$. In (39) necessarily $a^\dagger \equiv a \pmod 2$; and the number of values of $a^\dagger$ to be considered may be diminished by employing (31), which cannot affect (39), since $\bar{u}a^\dagger u = \bar{v}a^\ddagger v$, *if* $v = iu$ and $a^\ddagger = -\hat{\imath}a^\dagger i$, etc. All cases are virtually contained then in

LEMMA 1. *If $Na \equiv 3 \pmod 8$, and $a' = a$ or is obtained from a by permuting the a_i, then aAb if $b \equiv a' \pmod 4$, aBb if $b \equiv -a' \pmod 4$. If $Na \not\equiv 3 \pmod 8$, let a' be obtained from a by interchanging the coördinates a_f and a_g of like parity; and a'' from a' by changing the sign of the remaining coördinate. Then:*

$$
\begin{aligned}
&\text{if } Na \equiv 1 \pmod 8,\ aAb \text{ when } b \equiv \pm a,\ \pm a',\ \text{or } \pm a'' && \pmod 4;\\
&\text{if } Na \equiv 5 \pmod 8,\ aAb \text{ when } b \equiv \pm a && \pmod 4,\\
&\qquad\qquad\quad aBb \text{ when } b \equiv \pm a' \text{ or } \pm a'' && \pmod 4;\\
&\text{if } Na \equiv 2 \pmod{16},\ aAb \text{ when } b \equiv \pm a,\ \pm a',\ \text{or } \pm a'' && \pmod 8,\\
&\qquad\qquad\quad aBb \text{ when } b \equiv \pm 3a,\ \pm 3a',\ \text{or } \pm 3a'' && \pmod 8;\\
&\text{if } Na \equiv 6 \pmod{16},\ aAb \text{ when } b \equiv a,\ 3a,\ a',\ 3a',\ 5a'',\ 7a'' && \pmod 8,\\
&\qquad\qquad\quad aBb \text{ when } b \equiv 5a,\ 7a,\ 5a',\ 7a',\ a'',\ 3a'' && \pmod 8;\\
&\text{if } Na \equiv 10 \pmod{16},\ aAb \text{ when } b \equiv a,\ 7a,\ 3a',\ 5a',\ 3a'',\ 5a'' && \pmod 8,\\
&\qquad\qquad\quad aBb \text{ when } b \equiv 3a,\ 5a,\ a',\ 7a',\ a'',\ 7a'' && \pmod 8;\\
&\text{if } Na \equiv 14 \pmod{16},\ aAb \text{ when } b \equiv a,\ 3a,\ 5a',\ 7a',\ a'',\ 3a'' && \pmod 8,\\
&\qquad\qquad\quad aBb \text{ when } b \equiv 5a,\ 7a,\ a',\ 3a',\ 5a'',\ 7a'' && \pmod 8.
\end{aligned}
$$

We obtain next certain relations connected with

$$(40)\qquad at = -x_0 + ix_1 + jx_2 + kx_3 = \xi,\quad \check{\imath}at = ic_1 + jc_2 + kc_3 = c,$$

where, on expanding, we find

$$
(41)\qquad
\begin{aligned}
x_0 &= a_1t_1 + a_2t_2 + a_3t_3, & x_1 &= a_1t_0 + a_2t_3 - a_3t_2,\\
x_2 &= a_2t_0 + a_3t_1 - a_1t_3, & x_3 &= a_3t_0 + a_1t_2 - a_2t_1;
\end{aligned}
$$

$$
(42)\qquad
\begin{aligned}
c_1 &= a_1(t_0^2 + t_1^2 - t_2^2 + t_3^2) + 2a_2(t_0t_3 + t_1t_2) + 2a_3(-t_0t_2 + t_1t_3),\\
c_2 &= 2a_1(-t_0t_3 + t_1t_2) + a_2(t_0^2 + t_2^2 - t_1^2 - t_3^2) + 2a_3(t_0t_1 + t_2t_3),\\
c_3 &= 2a_1(t_0t_2 + t_1t_3) + 2a_2(-t_0t_1 + t_2t_3) + a_3(t_0^2 + t_3^2 - t_1^2 - t_2^2).
\end{aligned}
$$

We can easily verify that, to modulus h $(= \Sigma a_i^2)$,

$$
(43)\qquad
\begin{aligned}
a_1c_1 &\equiv x_0^2 + x_1^2, & a_1c_2 &\equiv x_0x_3 + x_1x_2, & a_1c_3 &\equiv -x_0x_2 + x_1x_3,\\
a_2c_1 &\equiv -x_0x_3 + x_1x_2, & a_2c_2 &\equiv x_0^2 + x_2^2, & a_2c_3 &\equiv x_0x_1 + x_2x_3,\\
a_3c_1 &\equiv x_0x_2 + x_1x_3, & a_3c_2 &\equiv -x_0x_1 + x_2x_3, & a_3c_3 &\equiv x_0^2 + x_3^2;
\end{aligned}
$$

$$(44)\qquad \Sigma x_t^2 = N\xi = Na \cdot Nt \equiv 0 \pmod h.$$

Similarly we write

$$(45)\qquad a^\dagger u = -y_0 + iy_1 + jy_2 + ky_3 = \eta,\qquad \bar{u}a^\dagger u = id_1 + jd_2 + kd_3 = d.$$

The substitutions which carry a into $a^\dagger$ replace (41)-(44) by (41†)-(44†) with $t \to u$, $x \to y$, $c \to d$, $\xi \to \eta$.

In order that $\check{\imath}at$ be prime to H, at must be likewise, i. e.

$$(46)\qquad x_0, x_1, x_2, x_3, H \text{ must be coprime.}$$

The hypothesis of (39) breaks up into two parts:

$$c_i \equiv \lambda d_i \pmod{H} \qquad (i = 1, 2, 3), \tag{47}$$

$$u\mathfrak{l}at\bar{u} \equiv \lambda a^\dagger \pmod{4l}. \tag{47'}$$

We shall show in all cases that, if (47) and (46) both hold, one of the four characters $(\pm \lambda|H)$, $(\pm 2\lambda|H)$ has a certain value, and then verify from Lemma 1 that in accordance with this value and (47'),

$$aA\lambda a^\dagger \text{ if } (lH|\lambda) = 1, \qquad aB\lambda a^\dagger \text{ if } (lH|\lambda) = -1. \tag{48}$$

First take $a^\dagger = a$. Now $a_i c_i \equiv \lambda a_i d_i \pmod{H}$, and by (43), (43†),

$$x_0^2 + x_i^2 \equiv \lambda(y_0^2 + y_i^2) \pmod{H} \qquad (i = 1, 2, 3). \tag{49}$$

Summing, we have $2x_0^2 + N\xi \equiv \lambda(2y_0^2 + N\eta)$, $x_0^2 \equiv \lambda y_0^2$,

$$x_f^2 \equiv \lambda y_f^2 \pmod{H} \qquad (f = 0, 1, 2, 3). \tag{50}$$

This contradicts (46) unless $(\lambda|H) = 1$, whence

$$\begin{aligned} (lH|\lambda) &= -(l|\lambda) \text{ if } H \equiv \lambda \equiv 3 \pmod 4, \\ &= (l|\lambda) \text{ otherwise.} \end{aligned} \tag{51}$$

Now $H \equiv h/l \pmod 8$. In Lemma 1, if $h \equiv 3 \pmod 8$, $aA\lambda a$ or $aB\lambda a$ according as $\lambda \equiv 1$ or $-1 \pmod 4$, in agreement with (51); if $H \equiv 3 \pmod 4$ and $l = 2$, $aA\lambda a$ if $\lambda \equiv 1$ or $3 \pmod 8$, $aB\lambda a$ if $\lambda \equiv 5$ or $7 \pmod 8$, which agrees with (51). If $H \not\equiv 3 \pmod 4$ a glance at Lemma 1 confirms that $aA\lambda a$ or $aB\lambda a$ according as $(l|\lambda) = 1$ or -1.

Second take $a^\dagger = ia_1 + ja_3 + ka_2$, assuming $a_2 \equiv a_3 \pmod 2$. This is the a' of Lemma 1. The cases $a_1 \equiv a_2$ or $a_3 \equiv a_1 \pmod 2$ will follow by symmetry.

Now by (43) and (43†), and (44†),

$$\begin{gathered} a_1c_1 \equiv -x_2^2 - x_3^2,\ a_2c_2 + a_3c_3 \equiv 2x_0^2 + x_2^2 + x_3^2,\ a_2c_3 + a_3c_2 \equiv 2x_2x_3, \\ a_1d_1 \equiv -y_2^2 - y_3^2,\ a_3d_3 + a_2d_2 \equiv 2y_2y_3,\ a_3d_2 + a_2d_3 \equiv 2y_0^2 + y_2^2 + y_3^2, \end{gathered}$$

and by (47),

$$\begin{aligned} &2x_0^2 \equiv -\lambda(y_2 - y_3)^2, \qquad 2\lambda y_0^2 \equiv -(x_2 - x_3)^2, \\ &(x_2 + x_3)^2 \equiv -2\lambda y_1^2, \qquad \lambda(y_2 + y_3)^2 \equiv -2x_1^2, \pmod H, \end{aligned} \tag{52}$$

whence, unless $(-2\lambda|H) = 1$, some prime factor of H divides x_0, x_1, $x_2 + x_3$, and $x_2 - x_3$, contrary to (46). Hence

$$\begin{aligned} (lH|\lambda) &= -(l|\lambda)(-2|H) \text{ if } H \equiv \lambda \equiv 3 \pmod 4, \\ &= (l|\lambda)(-2|H) \text{ otherwise.} \end{aligned} \tag{52'}$$

The comparison with (48) and Lemma 1 is left to the reader.

Third take $a^\dagger = -ia_1 + ja_3 + ka_2$, $a_2 \equiv a_3 \pmod 2$. This is equivalent, by (31), to $a^\dagger = -ia_1 - ja_3 - ka_2$, and hence may be reduced to the preceding case by changing λ to $-\lambda$. Thus we have to compare (48) with $a^\dagger = a''$, and Lemma 1, upon the assumption that $(2\lambda|H) = 1$, i. e.

$$\begin{aligned}(lH|\lambda) &= -(l|\lambda)(2|H) \text{ if } H \equiv \lambda \equiv 3 \pmod 4,\\ &= \ \ (l|\lambda)(2|H) \text{ otherwise.}\end{aligned}$$

The accordance with Lemma 1 is easily verified.

There now remains only $a_1 \equiv a_2 \equiv a_3 \pmod 2$, whence $Na \equiv 3 \pmod 8$; with $a^\dagger$ equal to α) $ia_2 + ja_3 + ka_1$, β) $-ia_2 - ja_3 - ka_1$, γ) $ia_3 + ja_1 + ka_2$, or δ) $-ia_3 - ja_1 - ka_2$.

Fourth, assume α). By (43) and (43†),

$$a_1c_1 \equiv x_0^2 + x_1^2 \equiv -x_2^2 - x_3^2,\ a_2c_3 \equiv x_0x_1 + x_2x_3,\ a_3c_2 \equiv -x_0x_1 + x_2x_3,$$
$$a_1d_1 \equiv y_0y_2 + y_1y_3,\ a_2d_3 \equiv -y_0y_2 + y_1y_3,\ a_3d_2 \equiv y_0^2 + y_2^2 \equiv -y_1^2 - y_3^2,$$

whence by (47),

$$(x_0 + x_1)^2 \equiv \lambda(y_1 + y_3)^2, \qquad (x_0 - x_1)^2 \equiv \lambda(y_0 + y_2)^2,$$
$$(x_2 + x_3)^2 \equiv \lambda(y_0 - y_2)^2, \qquad (x_2 - x_3)^2 \equiv \lambda(y_1 - y_3)^2,$$

contradicting (46) unless $(\lambda|H) = 1$. That is, since $Na = h \equiv H \equiv 3 \pmod 8$, $(H|\lambda) = (-1|\lambda)$. This agrees with the first part of Lemma 1 and (48).

Case β) is reduced to α) by changing λ to $-\lambda$; hence $(-\lambda|H) = 1$ and $(H|\lambda) = -(-1|\lambda)$, in agreement with (48) and Lemma 1 with $Na \equiv 3 \pmod 8$. Cases γ) and δ) are similar.

As a corollary we have

THEOREM 8. *Let H be an odd factor of Na, a as in (5), λ an integer. Then the congruence*

$$\bar{\imath}at \equiv \lambda\bar{u}a^\dagger u \pmod H, \tag{53}$$

has no solutions t, u with $\bar{\imath}at$ proper, if

1) *λ is a quadratic non-residue of H, and (apart from operations (31)), $a^\dagger = a$ or $ia_2 + ja_3 + ka_1$ or $ia_3 + ja_1 + ka_2$;* (53′)

or

2) *-2λ is a quadratic non-residue of H, and (apart from operations (31)), $a^\dagger = ia_1 + ja_3 + ka_2$ or $ia_3 + ja_2 + ka_1$ or $ia_2 + ja_1 + ka_3$.* (53″)

By changing λ into $-\lambda$, $a^\dagger$ into $-a^\dagger$, we have a result associated with any odd or even number of permutations or of sign-changes of the a_i.

THEOREM 9. *Let H be any odd integer, λ an integer prime to H, a proper*

and pure. Then (53) *is solvable for* t, u *with* $\bar{\imath}at$ *proper, if* (53′) *holds and* λ *is a quadratic residue of* H, *or if* (53″) *holds and* -2λ *is a quadratic residue of* H.

For let u be any quaternion for which $\bar{u}a^\dagger u$ is proper (mod H), e. g. $u = 1$. If $\lambda \equiv s^2 \pmod{H}$, and $a^\dagger = a$, take $t \equiv su$; if however $a^\dagger = ia_2 + ja_3 + ka_1$, take $t \equiv su(1 + i + j + k)/2 \pmod{H}$; and if $a^\dagger = ia_3 + ja_1 + ka_2$, take $t \equiv su(1 - i - j - k)/2$. The trivial effect of operations (31) is by now obvious. If $\lambda \equiv -2s^2 \pmod{H}$, and $a^\dagger = ia_1 - ja_3 - ka_2$, take $t \equiv s(j + k)$; and so forth.

Taking $\lambda = 1$ we have at once the statement referred to § 11 on p. 895.

12. THEOREM 10. *Let* a, b *be pure integral quaternions,*

$$p \text{ an odd prime, } p \nmid a,\ p \nmid b,\ p^r \mid Na,\ p^r \mid Nb,\ r > 0. \tag{54}$$

Then the number of solutions $t \pmod{p^r}$ *of*

$$\bar{\imath}at \equiv b \pmod{p^r} \tag{55}$$

is exactly p^{2r-2} *times the number of solutions* $t \pmod{p}$ of

$$\bar{\imath}at \equiv b \pmod{p}. \tag{56}$$

Clearly, any solution t of (55) must satisfy

$$\bar{\imath}at \equiv b \pmod{p^{r-1}}. \tag{57}$$

The solutions of (55) are to be sought among the quaternions $t + up^{r-1}$, where t ranges over all solutions of (57), and u is to be determined (mod p) to satisfy

$$(\bar{t} + p^{r-1}\bar{u})a(t + p^{r-1}u) \equiv b \pmod{p^r}.$$

On setting $\bar{\imath}at - b = p^{r-1}v$, this reduces to

$$v + \bar{u}at - \bar{\imath}au \equiv 0 \pmod{p},$$

and hence on writing $at = -x_0 + \cdots$ as in (40), we have

$$\begin{aligned} x_1u_0 + x_0u_1 - x_3u_2 + x_2u_3 &\equiv -\tfrac{1}{2}v_1, \\ x_2u_0 + x_3u_1 + x_0u_2 - x_1u_3 &\equiv -\tfrac{1}{2}v_2, \\ x_3u_0 - x_2u_1 + x_1u_2 + x_0u_3 &\equiv -\tfrac{1}{2}v_3, \pmod{p}. \end{aligned} \tag{58}$$

These three congruences are not independent. For if $c = \bar{\imath}at$,

$$\xi c = at\bar{\imath}at = -Nt(Na)t \equiv 0,$$

whence

$$(59)\qquad \begin{aligned} &c_1x_1 + c_2x_2 + c_3x_3 \equiv 0, \qquad c_1x_0 + c_2x_3 - c_3x_2 \equiv 0,\\ &-c_1x_3 + c_2x_0 + c_3x_1 \equiv 0, \qquad c_1x_2 - c_2x_1 + c_3x_0 \equiv 0, \pmod{p}. \end{aligned}$$

Fortunately, the same relation connects the right members of (58):

$$(60)\qquad c_1v_1 + c_2v_2 + c_3v_3 \equiv 0 \pmod{p}.$$

To prove this we consider $c - b = p^{r-1}v$,

$$(\bar{c} - \bar{b})(c - b) = p^{2r-2}\bar{v}v,$$

whence as $p^r \mid Nc$ and $p^r \mid Nb$, $p^r \mid \frac{1}{2}(\bar{c}b + \bar{b}c) =$ real part of $\bar{c}b$. But $\bar{c}c - \bar{c}b = p^{r-1}\bar{c}v$. Hence p divides the real part of cv, which is (60).

Thus the congruences (58) reduce to two. For example let $p \nmid c_3$. Then (58_3) is a combination of (58_1) and (58_2). By (43), one of the determinants $-x_0x_2 + x_1x_3$, $x_0x_1 + x_2x_3$, $x_0^2 + x_3^2$ is prime to p. Hence we can choose two of u_0, u_1, u_2, u_3 arbitrary (mod p), and the remaining two are then uniquely determined. That is, (58) has precisely p^2 solutions u (mod p), and (55) has exactly p^2 times the number of solutions of (57).

13. THEOREM 11. *Let a be a pure quaternion, $p \mid Na$, $p \nmid a$, p an odd prime. Then*

$$(61)\qquad \bar{\imath}at \equiv 0 \pmod{p} \text{ for exactly } p^2 \text{ residues } t \pmod{p};$$

and $\bar{\imath}at$ represents (mod p) exactly $\frac{1}{2}(p^2 - 1)$ of the $p^2 - 1$ residues b such that $p \nmid b$ and $p \mid Nb$, and each such residue represented is obtained for exactly $2p^2$ incongruent values t (mod p).

We inquire, for a given u, how many values t (mod p) satisfy

$$(62)\qquad \bar{\imath}at \equiv \bar{u}au \pmod{p}.$$

We employ notations (40)-(45), with $a^\dagger = a$. Hence (62) becomes $c_i \equiv d_i$ (mod p) $(i = 1, 2, 3)$.

Thus (62) implies $a_ic_i \equiv a_id_i$, and as in (50),

$$(63)\qquad x_f^2 \equiv y_f^2 \pmod{p}, \qquad (f = 0, 1, 2, 3).$$

Conversely either of the systems

$$(64)\qquad x_f \equiv y_f \pmod{p}, \qquad (f = 0, 1, 2, 3),$$

$$(65)\qquad x_f \equiv -y_f \pmod{p}, \qquad (f = 0, 1, 2, 3),$$

implies in view of (43) and $(43^\dagger)$, $a_ic_j \equiv a_id_j$ $(i, j = 1, 2, 3)$, and hence, since $p \nmid a$, $c_j \equiv d_j$ $(j = 1, 2, 3)$, which is (62).

Now (64) and (65) each possess exactly p^2 incongruent solutions $t \pmod p$. To see this we write out (64) in full:

$$(66)\quad \begin{array}{rrrrl} a_1t_1 + a_2t_2 + a_3t_3 \equiv & & & a_1u_1 + a_2u_2 + a_3u_3, \\ a_1t_0 & - a_3t_2 + a_2t_3 \equiv & a_1u_0 & - a_3u_2 + a_2u_3, \\ a_2t_0 + a_3t_1 & - a_1t_3 \equiv & a_2u_0 + a_3u_1 & - a_1u_3, \\ a_3t_0 - a_2t_1 + a_1t_2 & \equiv & a_3u_0 - a_2u_1 + a_1u_2, & \pmod p. \end{array}$$

The argument for $p \nmid a_1$ will be typical. Then (66_3), (66_4) are linear combinations of (66_1) and (66_2), with the multipliers a_3/a_1 and a_2/a_1, $-a_2/a_1$ and a_3/a_1. The system reduces to the first two congruences; t_2 and t_3 can be chosen arbitrarily, and t_0, t_1 are then uniquely determined; that is, t has p^2 residues $\pmod p$. The same argument applies to (65).

The solutions t of (64) are distinct from those of (65) unless $y_f \equiv -y_f \equiv 0$ $(f = 0, 1, 2, 3)$. Then $au \equiv 0$, $\bar{u}au \equiv 0$. The case $\bar{t}at \equiv 0 \pmod p$ can now be completed. We can take $u = 0$, and see that (66) with zeros on the right has again exactly p^2 solutions. This yields (61). Incidentally it shows that

$$(67)\qquad \text{if } p \nmid a,\ p \mid Na, \text{ then } p \mid \bar{t}at \text{ if and only if } p \mid at.$$

For the rest we shall suppose $p \nmid c = \bar{t}at$. In addition to (59) we must, if $c_i \equiv d_i$, have

$$(68)\quad \begin{array}{ll} c_1y_1 + c_2y_2 + c_3y_3 \equiv 0, & c_1y_0 + c_2y_3 - c_3y_2 \equiv 0, \\ -c_1y_3 + c_2y_0 + c_3y_1 \equiv 0, & c_1y_2 - c_2y_1 + c_3y_0 \equiv 0, \pmod p. \end{array}$$

There remains from (63) to be treated the possibilities $x_f \equiv \pm y_f$ in which not all the signs are alike. We shall reduce all cases to (64) or (65).

Suppose for example, $x_0 \equiv -y_0$, $x_1 \equiv y_1$, $x_2 \equiv y_2$, $x_3 \equiv y_3$. From (59) and (68) follow $p \mid 2c_iy_0$ $(i = 1, 2, 3)$, $p \mid y_0$. Hence $x_0 \equiv y_0$ and we have (64). Similarly whenever one or three of the signs are $+$ (or $-$) we are led to (64) or (65).

The case of two plus or two minus signs is typified by

$$(69)\qquad x_0 \equiv -y_0,\ x_1 \equiv -y_1,\ x_2 \equiv +y_2,\ x_3 \equiv +y_3.$$

If $c_i \equiv d_i$, then (59) and (68) yield

$$p \mid c_1,\ p \mid c_2y_2 + c_3y_3,\ p \mid c_2y_0 + c_3y_1,$$

and since $p \nmid c$, $p \mid c_2^2 + c_3^2$ but $p \nmid c_2c_3$. Incidentally, $p \equiv 1 \pmod 4$. Further by (43),

$$a_1c_2 \equiv x_0x_3 + x_1x_2 \equiv -y_0y_3 - y_1y_2 \equiv -a_1d_2,$$

whence $p|a_1$ also (since $c_2 \equiv d_2 \not\equiv 0$), and $p \nmid a_2a_3$. Thus

$$x_0 \equiv a_2t_2 + a_3t_3 \equiv (-a_2/a_3)(-a_3t_2 + a_2t_3) \equiv (-a_2/a_3)x_1,$$

and similarly $x_2 \equiv (a_2/a_3)x_3$. Thus by (42_1) and $p|c_1$,

$$\begin{aligned} 0 \equiv c_1 &\equiv 2t_0(a_2t_3 - a_3t_2) + 2t_1(a_2t_2 + a_3t_3) \\ &\equiv 2t_0x_1 - 2t_1(a_2/a_3)x_1 \equiv (2/a_3)x_1(a_3t_0 - a_2t_1) \equiv 2x_1x_3/a_3, \end{aligned}$$

that is either $x_0 \equiv x_1 \equiv 0$, or $x_2 \equiv x_3 \equiv 0$; reducing the case to (64) or (65).

The theorem follows, since t has p^4 possible residues (mod p), p^2 of these are used for $\bar{\imath}at \equiv 0$, and the remaining $p^2(p^2-1)$ form sets of $2p^2$ residues t for each of which $\bar{\imath}at$ has the same residue. That

$$b_1^2 + b_2^2 + b_3^2 \equiv 0 \pmod{p} \tag{70}$$

has exactly p^2 solutions, including $(0, 0, 0)$, is easy to show.

The characterization by other criteria of the residues $\bar{\imath}at \pmod{p}$ is not here attempted.

COROLLARY 1. *Congruence* (56) *is solvable for* $b \equiv 0$ *and for only half those residues* $b \not\equiv 0 \pmod{p}$ *such that* $Nb \equiv 0$.

COROLLARY 2. *If* $p|Na$, $p > 2$, *then* $\bar{\imath}at \equiv \bar{\imath}au$ *implies* $at \equiv au$ *or* $at \equiv -au \pmod{p}$.

14. Write $a^* = ia_1 + ja_3 + ka_2$, and assume $gcd(a_1, a_2, a_3) = 1$. By Theorem 8 with $\lambda = \pm 1$, and Theorem 11 *et seq.*, we see that if H is an odd factor of Na, then all residues $b \pmod{H}$, such that $Nb \equiv 0$, are represented between

$$\begin{aligned} &\bar{\imath}at \text{ and } \bar{\imath}(-a)t, && \text{when } H = p \equiv 3 \pmod 4, \\ &\bar{\imath}at \text{ and } \bar{\imath}a^*t, && \text{when } H = p \equiv 5 \text{ or } 7 \pmod 8, \\ &\bar{\imath}at \text{ and } \bar{\imath}(-a^*)t, && \text{when } H = p \equiv 3 \text{ or } 5 \pmod 8, \\ &\pm\bar{\imath}at \text{ and } \pm\bar{\imath}a^*t, && \text{when } H = pq,\ p,\ q, \text{ and } 1 \text{ incongruent } \pmod 8; \end{aligned} \tag{71}$$

p and q representing odd primes.

It will be observed that if, in Theorem 1, h' has no odd factor > 1 of the form $8f + 1$, it is of the form lH, with

$$H = 1 \text{ or } p \text{ or } pq, \tag{72}$$

p, q odd primes incongruent to each other and to 1 (mod 8).

Evidently all solutions of (1) satisfy

$$Nx \equiv lH \pmod{8lH}, \tag{73}$$

if H is an odd divisor of h such that $h/(lH) \equiv 1 \pmod 8$. Further, if H is of type (72), then as a ranges over a set $\mathfrak{S}(a)$, and t over all integral quaternions, $\bar{\imath}at$ represents *all* residues $x \pmod{4lH}$ satisfying (73). For by Theorem 4, since $Na = h \equiv lH \pmod{8l}$, every residue $x \pmod{4l}$ such that

$$x \equiv a \pmod 2, \qquad Nx \equiv lH \pmod{8l} \tag{74}$$

is represented by each of $\bar{\imath}at$ and $-\bar{\imath}at$ (separately), and every residue $x \equiv a^* \pmod 2$ satisfying (74_2) by each of $\bar{\imath}a^*t$ and $-\bar{\imath}a^*t$. We can choose a so that $a_2 \equiv a_3 \pmod 2$, $a \equiv a^* \pmod 2$, and hence can combine all residues $(\mathrm{mod}\ H)$ with all residues $(\mathrm{mod}\ 4l)$ such that $x \equiv a \pmod 2$. Permuting a_1, a_2, a_3 cyclically we get all residues $x \pmod{4lH}$ subject to (73). This proves the first part of II in Theorem 1.

15. To prove the second part of II in Theorem 1, consider a solution x of (1) with $(h/h')(8n+1)$ not a square. We can choose an odd prime p such that simultaneously

$$(-h' \mid p) = -1, \qquad (-h(8n+1) \mid p) = 1. \tag{75}$$

Choose s so that $s^2 + h(8n+1) \equiv 0 \pmod p$. Then $s + x$ has exactly eight (left-associate) right divisors of norm p, say

$$s + x = uv, \qquad Nv = p. \tag{76}$$

Then

$$vx\bar{v} = px', \text{ where } x' = vu - s. \tag{77}$$

The relation between x and x' is reciprocal, $\bar{v}$ carrying x' back into x. If v is replaced by a left-associate, the only change in x' is that two components are changed in sign. Thus a $(1, 1)$ correspondence is set up connecting the four solutions

$$(x_1, x_2, x_3),\ (x_1, -x_2, -x_3),\ (-x_1, x_2, -x_3),\ (-x_1, -x_2, x_3),$$

of (1), with four similarly related solutions x', $-ix'i$, etc.

We shall prove that x and x' are in opposite classes A and B, in view of (75_1), provided $h' = lH$, H of form (72). Suppose $x \equiv \bar{\imath}at \pmod{4h'}$, which is possible by § 14; and let $x' \equiv \bar{w}a^\dagger w \pmod{4h'}$, where $a^\dagger$ is also in $\mathfrak{S}(a)$. Then by (77),

$$v\bar{\imath}at\bar{v} \equiv p\bar{w}a^{\dagger}w \pmod{4h'}.$$

By Theorem 7,

$$N(vt) \equiv Nw \pmod 4 \text{ if } (h'|p) = 1,$$
$$\equiv -Nw \text{ if } (h'|p) = -1.$$

If $(h'|p) = 1$, $p \equiv 3 \pmod 4$ by (75_1), $Nt \equiv -Nw \pmod 4$, x' is in the opposite class to x. If $(h'|p) = -1$, $p \equiv 1 \pmod 4$, and again $Nv \equiv -Nw$.

16. In Theorem 1 let $h' = lH$ have no factor > 1 and $\equiv 1 \pmod 8$. By § 14 every solution x of (73) with x_1, x_2, x_3, h' coprime, belongs to A or B. By Theorem 7,

(78) if λ is an integer prime to $4h'$, then x and λx are in the same or opposite classes A or B, according as $(h'|\lambda) = 1$ or -1.

For example, in the table following Theorem 1, if $h' = 2$ and $\lambda = 3$, $(2|3) = -1$, $(1, 1, 0)$ is in A, $(3, 3, 0)$ in B.

17. If $h = H$ or $2H$, H as in (72), then for any solution x of (1) there exist integers r_1, r_2, r_3, and an odd integer m such that

$$hm^2 = (x_1 + 4r_1h)^2 + (x_2 + 4r_2h)^2 + (x_3 + 4r_3h)^2. \tag{79}$$

For we can solve $x \equiv \bar{\imath}at \pmod{4h}$ with $Na = h$ and Nt odd.

McGill University.

A LIST OF EXPANSIONS FOR THE FUNCTION $\phi(x, y, z)$.*

By WALTER H. GAGE.

1. In a recent paper [1] the writer has given a method for obtaining the trigonometric expansions for the so-called doubly periodic functions $\phi_{abcd}(x, y, z)$, where

$$\phi_{abcd}(x, y, z) \equiv \frac{\vartheta_1'^2 \vartheta_a(x+y+z)}{\vartheta_b(x)\vartheta_c(y)\vartheta_d(z)},$$

for sixty-four values of the quadruple index *abcd*.

Using the set of sixteen expansions for the so-called doubly periodic functions $\phi_{abc}(x, y)$ of the second kind [2] (first obtained by Hermite) and the relation

$$\phi_{abcd}(x, y, z) = \phi_{ebc}(x, y) \cdot \phi_{aed}(x+y, z),$$

we obtain expansions which are valid where the regions of convergence of the functions $\phi_{ebc}(x, y)$, $\phi_{aed}(x+y, z)$ overlap.[3] By this method it is possible to get the expansions for *abcd* equal to 0000, 1111, 2222, 3333, 0011, 0022, 0033, 1122, 1133, 2233, 0123, and all possible arrangements of each of these numbers. In the lists below, however, we have omitted those which are obtainable from the given ones merely by interchanging x, y, z.

The two sets of expansions obtained by Basoco and Bell [4] for forty-eight functions in two variables prove to be special cases of these new expansions for x or y equal to zero. The case $z = 0$ gives us for the forty-eight functions a third set which so far has not been published. On the other hand, the method employed by Basoco and Bell can be applied to the sixty-four $\phi_{abcd}(x, y, z)$ to yield four sets of expansions for one hundred and ninety-two additional functions in three variables.

---

* Received June 17, 1937.

[1] "A method of obtaining certain theta expansions," to appear in the *Transactions of the Royal Society of Canada.*

[2] E. T. Bell, "Arithmetical paraphrases," *Transactions of the American Mathematical Society*, vol. 22 (1921), pp. 198-219; "Theta expansions useful in arithmetic," *Messenger of Mathematics*, no. 635, vol. 54 (1924), pp. 166-176.

[3] Tannery and Molk, "Éléments de la Théorie des Fonctions Elliptiques," Tome 3, pp. 120-129.

[4] "Further theta expansions useful in arithmetic," *American Journal of Mathematics*, vol. 57 (1935), pp. 1-10.

2. As in the case of the series in two variables, these new expansions enable us to derive by the method of paraphrase a large body of arithmetical theorems on quadratic forms. In particular a full set, applied systematically to Jacobi's identities [5] (and to all which can be derived from them) expressing $\vartheta_\alpha(x+y+z)$ in terms of $\vartheta_\beta(x)$, $\vartheta_\gamma(y)$, $\vartheta_\delta(z)$, leads to an endless variety of paraphrases, involving arbitrary odd or even functions of divisors of integers in given partitions. In the general case, relations between ϑ and ϕ functions enable us to get paraphrases in any even number of variables. As a matter of fact, since certain of the $\phi_{abcd}(x, y, z)$ possess alternate expansions, direct comparison of the two series in these cases gives at once a set of arithmetical theorems. A study of the results that have been obtained by using the expansions for $\phi_{abc}(x, y)$ gives ample evidence of the possible applications of these new expansions in deriving theorems on quadratic partitions.

In what follows the reader is referred to Bell's papers cited above,[2] and in particular to the paper by Basoco and Bell,[4] for explanations of notation.

3. Following their method we set down the various partitions which occur.

(I) $\quad 2n = 2n_1 + 2n_2, \quad n = d\delta, \quad n_1 = d_1\delta_1, \quad n_2 = d_2\delta_2.$

(II) $\quad n = n_1 + n_2, \quad n = t\tau, \quad n_1 = t_1\tau_1, \quad n_2 = t_2\tau_2.$

(III) $\quad m = 4n_1 + m_2, \quad m = t\tau, \quad n_1 = d_1\delta_1, \quad m_2 = t_2\tau_2.$

(IV) $\quad \begin{cases} m = n_1 + 2n_2, \quad m = t\tau, \quad n_1 = t_1\tau_1, \quad n_2 = d_2\delta_2. \\ 2n = n'_1 + 2n'_2, \quad 2n = t'\tau', \quad n = d\delta, \quad n'_1 = t'_1\tau'_1, \quad n'_2 = d'_2\delta'_2. \end{cases}$

(V) $\quad m = m_1 + 2n_2, \quad m = t\tau, \quad m_1 = t_1\tau_1, \quad n_2 = t_2\tau_2.$

4. The expansions of which the partition is (I) (in § 3) have the form

$$\phi_{abcd}(x, y, z) = T_{abcd}(x, y, z) + 4\sum_{(n)} q^{2n}(\sum F_{abcd}(x, y, z; n)).$$

If we let

$$P_1(x, y, z) = \cos 2\{(d_1 - d_2)x + (d_1 - \delta_2)y + \delta_1 z\} - \cos 2\{(d_1 + d_2)x + (d_1 + \delta_2)y + \delta_1 z\}$$

then

$$T_{1111}(x, y, z) = -1 + \cot x \cot y + \cot y \cot z + \cot z \cot x;$$
$$T_{2211}(x, y, z) = -1 - \tan x \cot y + \cot y \cot z - \cot z \tan x;$$
$$T_{1221}(x, y, z) = 1 - \tan x \tan y + \tan y \cot z + \cot z \tan x;$$
$$T_{2112}(x, y, z) = T_{2211}(z, y, x); \quad T_{1122}(x, y, z) = T_{1221}(z, y, x);$$
$$T_{2222}(x, y, z) = 1 - \tan x \tan y - \tan y \tan z - \tan z \tan x;$$

[5] Jacobi, *Werke*, vol. 1, p. 501 etc.

$$
\begin{aligned}
F_{1111}(x,y,z;\,n) &= F_{1111}(x,y,z) = \{\cot(x+y)+\cot z\}\sin 2(dx+\delta y)\\
&\quad+(\cot x+\cot y)\sin 2\{d(x+y)+\delta z\}+2P_1(x,y,z);\\
F_{2211}(x,y,z) &= (-1)^d[\{\cot z-\tan(x+y)\}\sin 2(dx+\delta y)\\
&\quad+(\cot y-\tan x)\sin 2\{d(x+y)+\delta z\}]\\
&\quad+2(-1)^{d_1+d_2}P_1(x,y,z);\\
F_{1221}(x,y,z) &= -(-1)^{d+\delta}\{\cot z+\cot(x+y)\}\sin 2(dx+\delta y)\\
&\quad+(\tan x+\tan y)\sin 2\{d(x+y)+\delta z\}\\
&\quad-2(-1)^{d_2+\delta_2}P_1(x,y,z);\\
F_{2112}(x,y,z) &= \{\cot(x+y)-\tan z\}\sin 2(dx+\delta y)\\
&\quad+(-1)^{\delta}(\cot x+\cot y)\sin 2\{d(x+y)+\delta z\}\\
&\quad+2(-1)^{\delta_1}P_1(x,y,z);\\
F_{1122}(x,y,z) &= (-1)^{\delta}[\{\tan(x+y)+\tan z\}\sin 2(dx+\delta y)\\
&\quad-(-1)^{d}(\cot x-\tan y)\sin 2\{d(x+y)+\delta z\}]\\
&\quad-2(-1)^{d_1+\delta_1+\delta_2}P_1(x,y,z);\\
F_{2222}(x,y,z) &= (-1)^{\delta}[(-1)^{d}\{\tan z-\cot(x+y)\}\sin 2(dx+\delta y)\\
&\quad+(\tan x+\tan y)\sin 2\{d(x+y)+\delta z\}]\\
&\quad-2(-1)^{d_2+\delta_1+\delta_2}P_1(x,y,z).
\end{aligned}
$$

5. The expansions in which the partition is (II) have the form $\phi_{abcd}(x,y,z) = T_{abcd}(x,y,z) + 4\sum_{(n)} q^n(\sum F_{abcd}(x,y,z;n))$.

Let

$$
\begin{aligned}
P_2(x,y,z) &= H_1(x,y,z), \quad R_2(x,y,z) = H_2(x,y,z),\\
Q_2(x,y,z) &= K_1(x,y,z), \quad S_2(x,y,z) = K_2(x,y,z),\\
H_k(x,y,z) &= \cos\{2(t_1-t_2)x+(2t_1-\tau_2)y+\tau_1 z\}\\
&\quad+(-1)^k\cos\{2(t_1+t_2)x+(2t_1+\tau_2)y+\tau_1 z\},\\
K_k(x,y,z) &= \sin\{2(t_1+t_2)x+(2t_1+\tau_2)y+\tau_1 z\}\\
&\quad-(-1)^k\sin\{2(t_1-t_2)x+(2t_1-\tau_2)y+\tau_1 z\}.
\end{aligned}
$$

We have

$$
\begin{aligned}
T_{0011}(x,y,z) &= T_{3311}(x,y,z) = \csc y\csc z;\\
T_{3021}(x,y,z) &= T_{0321}(x,y,z) = T_{3012}(x,z,y)\\
&= T_{0312}(x,z,y) = \sec y\csc z;\\
T_{0022}(x,y,z) &= T_{3322}(x,y,z) = \sec y\sec z;\\
F_{0011}(x,y,z) &= \csc y\sin\{2t(x+y)+\tau z\}\\
&\quad+\csc z\sin(2tx+\tau y)+2P_2(x,y,z);\\
F_{3311}(x,y,z) &= (-1)^n[\csc y\sin\{2t(x+y)+\tau z\}\\
&\quad+\csc z\sin(2tx+\tau y)+2P_2(x,y,z)];\\
F_{3021}(x,y,z) &= (-1)^n\sec y\sin\{2t(x+y)+\tau z\}\\
&\quad+(-1\,|\,\tau)\csc z\cos(2tx+\tau y)\\
&\quad+2(-1)^{n_1}(-1\,|\,\tau_2)Q_2(x,y,z);
\end{aligned}
$$

$$F_{0321}(x,y,z) = \sec y \sin\{2t(x+y) + \tau z\} + (-1)^n(-1|\tau)\csc z \cos(2tx + \tau y) + 2(-1)^{n_2}(-1|\tau_2)Q_2(x,y,z);$$

$$F_{3012}(x,y,z) = (-1|\tau)\csc y \cos\{2t(x+y) + \tau z\} + \sec z \sin(2tx + \tau y) + 2(-1|\tau_1)S_2(x,y,z);$$

$$F_{0312}(x,y,z) = (-1)^n[(-1|\tau)\csc y \cos\{2t(x+y) + \tau z\} + \sec z \sin(2tx + \tau y) + 2(-1|\tau_1)S_2(x,y,z)];$$

$$F_{0022}(x,y,z) = (-1|\tau)[(-1)^n \sec y \cos\{2t(x+y) + \tau z\} + \sec z \cos(2tx + \tau y)] + 2(-1)^{n_1}(-1|\tau_1\tau_2)R_2(x,y,z);$$

$$F_{3322}(x,y,z) = (-1|\tau)[\sec y \cos\{2t(x+y) + \tau z\} + (-1)^n \sec z \cos(2tx + \tau y)] + 2(-1)^{n_2}(-1|\tau_1\tau_2)R_2(x,y,z).$$

6. For partition (III) we have the form

$$\phi_{abcd}(x,y,z) = 4\sum_{(m)} q^{m/2}(\sum F_{abcd}(x,y,z;m)).$$

Let

$$P_3(x,y,z) = \cos\{(2d_1 - t_2)x + (2d_1 - \tau_2)y + 2\delta_1 z\} - \cos\{(2d_1 + t_2)x + (2d_1 + \tau_2)y + 2\delta_1 z\},$$

$$Q_3(x,y,z) = \sin\{(2d_1 + t_2)x + (2d_1 + \tau_2)y + 2\delta_1 z\} + \sin\{(2d_1 - t_2)x + (2d_1 - \tau_2)y + 2\delta_1 z\}.$$

Then

$$F_{1001}(x,y,z) = \{\cot(x+y) + \cot z\}\sin(tx + \tau y) + 2P_3(x,y,z);$$

$$F_{2301}(x,y,z) = (-1|t)\{\cot z - \tan(x+y)\}\cos(tx + \tau y) + 2(-1)^{d_1}(-1|t_2)Q_3(x,y,z);$$

$$F_{1331}(x,y,z) = (-1|m)\{\cot(x+y) + \cot z\}\sin(tx + \tau y) + (-1|m_2)P_3(x,y,z);$$

$$F_{2002}(x,y,z) = \{\cot(x+y) - \tan z\}\sin(tx + \tau y) + 2(-1)^{\delta_1}P_3(x,y,z);$$

$$F_{1032}(x,y,z) = (-1|\tau)\{\tan(x+y) + \tan z\}\cos(tx + \tau y) - 2(-1)^{d_1+\delta_1}(-1|\tau_2)Q_3(x,y,z);$$

$$F_{2332}(x,y,z) = (-1|m)\{\cot(x+y) - \tan z\}\sin(tx + \tau y) + 2(-1)^{\delta_1}(-1|m_2)P_3(x,y,z).$$

7. The form in which the partition is (IV) is

$$\phi_{abcd}(x,y,z) = T_{abcd}(x,y,z) + 4\sum_{(m)} q^m(\sum F_{abcd}(x,y,z;m)) + 4\sum_{(n)} q^{2n}(\sum G_{abcd}(x,y,z;n)).$$

We have

$$T_{0110}(x,y,z) = T_{3113}(x,y,z) = \csc x \csc y;$$
$$T_{0220}(x,y,z) = T_{3223}(x,y,z) = \sec x \sec y;$$
$$T_{3120}(x,y,z) = T_{0123}(x,y,z) = \csc x \sec y.$$

Let

$$P_4(x,y,z;t_1,\tau_1,d_2,\delta_2) = P_4(x,y,z),$$
$$P_4(x,y,z;t'_1,\tau'_1,d'_2,\delta'_2) = P'_4(x,y,z),$$
$$Q_4(x,y,z;t_1,\tau_1,d_2,\delta_2) = Q_4(x,y,z),$$
$$Q_4(x,y,z;t'_1,\tau'_1,d'_2,\delta'_2) = Q'_4(x,y,z);$$
$$P_4(x,y,z) = \cos\{(\tau_1 - 2d_2)x + (\tau_1 - 2\delta_2)y + 2t_1 z\} - \cos\{(\tau_1 + 2d_2)x + (\tau_1 + 2\delta_2)y + 2t_1 z\},$$
$$Q_4(x,y,z) = \sin\{(\tau_1 + 2d_2)x + (\tau_1 + 2\delta_2)y + 2t_1 z\} - \sin\{(\tau_1 - 2d_2)x + (\tau_1 - 2\delta_2)y + 2t_1 z\}.$$

Then

$$F_{0110}(x,y,z) = (\cot x + \cot y)\sin\{2tz + \tau(x+y)\} + 2P_4(x,y,z);$$
$$G_{0110}(x,y,z) = (\cot x + \cot y)\sin\{2t'z + \tau'(x+y)\} + \csc(x+y)\sin 2(dx+\delta y) + 2P'_4(x,y,z);$$
$$F_{3120}(x,y,z) = (-1|\tau)(\cot x - \tan y)\cos\{2tz + \tau(x+y)\} + 2(-1)^{\delta_2}(-1|\tau_1)Q_4(x,y,z);$$
$$G_{3120}(x,y,z) = (-1|\tau')(\cot x - \tan y)\cos\{2t'z + \tau'(x+y)\} + (-1)^{\delta}\sec(x+y)\sin 2(dx+\delta y) + 2(-1)^{\delta'_2}(-1|\tau'_1)Q'_4(x,y,z);$$
$$F_{0220}(x,y,z) = (\tan x + \tan y)\sin\{2tz + \tau(x+y)\} - 2(-1)^{d_2+\delta_2}P_4(x,y,z);$$
$$G_{0220}(x,y,z) = (\tan x + \tan y)\sin\{2t'z + \tau'(x+y)\} - (-1)^{d+\delta}\csc(x+y)\sin 2(dx+\delta y) - 2(-1)^{d'_2+\delta'_2}P'_4(x,y,z);$$
$$F_{3113}(x,y,z) = -(\cot x + \cot y)\sin\{2tz + \tau(x+y)\} - 2P_4(x,y,z);$$
$$G_{3113}(x,y,z) = (\cot x + \cot y)\sin\{2t'z + \tau'(x+y)\} + \csc(x+y)\sin 2(dx+\delta y) + 2P'_4(x,y,z);$$
$$F_{0123}(x,y,z) = -(-1|\tau)(\cot x - \tan y)\cos\{2tz + \tau(x+y)\} - 2(-1)^{\delta_2}(-1|\tau_1)Q_4(x,y,z);$$
$$G_{0123}(x,y,z) = (-1|\tau')(\cot x - \tan y)\cos\{2t'z + \tau'(x+y)\} + (-1)^{\delta}\sec(x+y)\sin 2(dx+\delta y) + 2(-1)^{\delta'_2}(-1|\tau'_1)Q'_4(x,y,z);$$

$$F_{3223}(x, y, z) = -(\tan x + \tan y)\sin\{2tz + \tau(x + y)\} + 2(-1)^{d_2+\delta_2}P_4(x, y, z);$$

$$G_{3223}(x, y, z) = (\tan x + \tan y)\sin\{2t'z + \tau'(x + y)\} - (-1)^{d+\delta}\csc(x + y)\sin 2(dx + \delta y) - 2(-1)^{d'_2+\delta'_2}P'_4(x, y, z).$$

8. Finally, those expansions in which the partition is (V) have the form

$$\phi_{abcd}(x, y, z) = 4\sum_{(m)} q^{m/2}(\sum F_{abcd}(x, y, z; m)).$$

On putting

$$P_5(x, y, z) = A_1(x, y, z), \quad Q_5(x, y, z) = A_2(x, y, z),$$
$$R_5(x, y, z) = B_1(x, y, z), \quad S_5(x, y, z) = B_2(x, y, z),$$
$$A_k(x, y, z) = \cos\{(t_1 - 2t_2)x + (t_1 - \tau_2)y + \tau_1 z\} + (-1)^k\cos\{(t_1 + 2t_2)x + (t_1 + \tau_2)y + \tau_1 z\},$$
$$B_k(x, y, z) = \sin\{(t_1 + 2t_2)x + (t_1 + \tau_2)y + \tau_1 z\} + (-1)^k\sin\{(t_1 - 2t_2)x + (t_1 - \tau_2)y + \tau_1 z\},$$

we have

$$F_{1010}(x, y, z) = \csc y\sin\{t(x + y) + \tau z\} + 2P_5(x, y, z);$$

$$F_{2310}(x, y, z) = (-1\,|\,t)\csc y\cos\{t(x + y) + \tau z\} + 2(-1)^{n_2}(-1\,|\,t_1)R_5(x, y, z);$$

$$F_{2020}(x, y, z) = (-1\,|\,t)\sec y\cos\{t(x + y) + \tau z\} + 2(-1\,|\,t_1\tau_2)Q_5(x, y, z);$$

$$F_{1320}(x, y, z) = \sec y\sin\{t(x + y) + \tau z\} + 2(-1)^{n_2}(-1\,|\,\tau_2)S_5(x, y, z);$$

$$F_{2013}(x, y, z) = (-1\,|\,\tau)\csc y\cos\{t(x + y) + \tau z\} + 2(-1\,|\,\tau_1)R_5(x, y, z);$$

$$F_{1313}(x, y, z) = (-1\,|\,m)\csc y\sin\{t(x + y) + \tau z\} + 2(-1)^{n_2}(-1\,|\,m_1)P_5(x, y, z);$$

$$F_{1023}(x, y, z) = (-1\,|\,m)\sec y\sin\{t(x + y) + \tau z\} + 2(-1\,|\,m_1)(-1\,|\,\tau_2)S_5(x, y, z);$$

$$F_{2323}(x, y, z) = (-1\,|\,\tau)\sec y\cos\{t(x + y) + \tau z\} + 2(-1)^{n_2}(-1\,|\,\tau_1\tau_2)Q_5(x, y, z).$$

Let

$$U_5(x, y, z) = C_1(x, y, z), \quad V_5(x, y, z) = C_2(x, y, z);$$
$$C_k(x, y, z) = \cos\{(\tau_2 - t_1)x + (\tau_2 - \tau_1)y + 2t_2 z\} + (-1)^k\cos\{(\tau_2 + t_1)x + (\tau_2 + \tau_1)y + 2t_2 z\}.$$

Then we have

$$F_{3030}(x,y,z) = (-1\,|\,\tau)\sec(x+y)\cos(tx+\tau y) + 2(-1\,|\,\tau_1\tau_2)V_5(x,y,z);$$

$$F_{0330}(x,y,z) = (-1\,|\,m)\csc(x+y)\sin(tx+\tau y) + 2(-1\,|\,m_1)U_5(x,y,z);$$

$$F_{3003}(x,y,z) = \csc(x+y)\sin(tx+\tau y) + 2(-1)^{n_2}U_5(x,y,z);$$

$$F_{3333}(x,y,z) = (-1\,|\,m)\csc(x+y)\sin(tx+\tau y) + 2(-1)^{n_2}(-1\,|\,m_1)U_5(x,y,z);$$

$$F_{0000}(x,y,z) = \csc(x+y)\sin(tx+\tau y) + 2U_5(x,y,z);$$

$$F_{0303}(x,y,z) = (-1\,|\,t)\sec(x+y)\cos(tx+\tau y) + 2(-1)^{n_2}(-1\,|\,t_1\tau_2)V_5(x,y,z).$$

UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, CANADA.

SOME ARITHMETICAL APPLICATIONS OF RESIDUATION.

By MORGAN WARD.

1°. The operation of residuation was apparently first considered by Dedekind in his theory of the modules in a ring of algebraic integers [1].[1] It was introduced into polynomial ideal theory by Emanuel Lasker [2], and has since been used systematically by F. S. Macaulay [3] and others. I propose to show here how the operation may be applied to various arithmetical problems,[2] in particular to developing a systematic calculus for the periods of elements in any finite Abelian group.

2°. Consider first for simplicity a cyclic group $\mathfrak{G}$ of order N written additively. Every element α of $\mathfrak{G}$ may be uniquely represented as

$$\alpha = a\gamma, \qquad 0 < a \leqq N \tag{2.1}$$

where γ is a fixed primitive element of $\mathfrak{G}$ and $a\gamma$ means $\gamma + \gamma + \cdots + \gamma$ taken a times. We write L_a for a in (2.1); for example $L_0 = N$. Let P_a denote the period of α; that is the least positive integer p such that $p\alpha = 0$. The starting point of our investigation is the observation that P_a *is the residual of* L_a *with respect to* N.

L_ξ considered as an operation on $\mathfrak{G}$ to the finite ring K_N of the integers modulo N is linear and distributive:

$$L_{ma+n\beta} \equiv mL_a + nL_\beta \pmod{N}, \qquad m, n \text{ integers.}$$

On the other hand, P_ξ is neither a linear nor a distributive operation; given P_a and P_β, all we can assert about $P_{a+\beta}$ is that it divides $[P_a, P_\beta]$, the least

[1] The numbers [1], [2], ... in square brackets refer to the bibliography at the close of the paper.

[2] For example, consider the problem of solving

$$AX \equiv 0 \ (\text{modd } \mathfrak{m}, F). \tag{1}$$

Here A and F are given polynomials in indeterminates $x_1, \cdots, x_s$ with coefficients in a commutative ring $\mathfrak{R}$ while $\mathfrak{m}$ is an ideal of $\mathfrak{R}$. We seek all solutions X in the quotient ring $\mathfrak{R}[x_1, \cdots, x_s]/\mathfrak{m}$. If $\mathfrak{A}$ and $\mathfrak{F}$ denote the ideals $(\mathfrak{m}, A)$, $(\mathfrak{m}, F)$ then the totality of such solutions of (1) constitute the residual ideal of $\mathfrak{A}$ with respect to $\mathfrak{F}$ (Van der Waerden [4] Chapter XII). Thus the solution of (1) is equivalent to specifying this residual, say by determining a basis for it. I have given a complete solution for the case when $s = 1$ and $\mathfrak{R}$ is the ring of rational integers. Ward [5].

common multiple of P_α and P_β.[3] Simple numerical examples show that $P_{\alpha+\beta}$ may be any divisor whatever of $[P_\alpha, P_\beta]$.

These facts suggest that we introduce in $\mathfrak{G}$ one or more new operations $\xi \circ \eta$, $\xi \times \eta$ such that $P_{\alpha\circ\beta}$, $P_{\alpha\times\beta}$ may be calculated knowing only the values of P_α, P_β; the definition of P_α as a residual immediately suggests how these operations should be defined. But before introducing these operations, we shall briefly summarize the properties of residuation of which we make use.

3. Let $\mathfrak{O}$ be the set of ideals[4] $A, B, C, \cdots$ of a fixed commutative ring containing a unit element. If A and B are any two elements of $\mathfrak{O}$, the residual of B with respect to A is by definition an ideal C such that

$$A \supset BC; \quad \text{if} \quad A \supset BX \quad \text{then} \quad C \supset X.$$

We write as usual $C = A : B$. The residual always exists and has the following properties:

$$\begin{aligned}
&(3.1) \quad A:B = A:(A,B) = [A,B]:B,\\
&\quad (A:B):C = (A:C):B = A:BC,\\
&\quad A = M:N \text{ and } B = M:(M:N) \text{ imply } B = M:A,\ A = M:B,\\
&\quad M:(A_1, A_2, \cdots, A_k) = [M:A_1, M:A_2, \cdots, M:A_k],\\
&\quad [A_1, A_2, \cdots, A_k]:M = [A_1:M, A_2:M, \cdots, A_k:M].
\end{aligned}$$

If we restrict $\mathfrak{O}$ to be a principal ideal ring, then $A \supset B$ if and only if there exists a quotient $Q = A/B$ such that $A = QB$. Furthermore this quotient is unique. It is easily shown that $A:B = A/B$ whenever the quotient A/B exists, so that formula (3.1) becomes

$$(3.11) \qquad A:B = \frac{A}{(A,B)} = \frac{B}{[A,B]}.$$

On using this result and the unicity of the quotient, we easily find that the following additional rules for residuation hold in any principal ideal ring.[5]

$$\begin{aligned}
&(M,N) = M:(M:N), \qquad M:AB = \{(M:A)(M:B)\}:M,\\
&\quad (A_1, A_2, \cdots, A_k):M = (A_1:M, \cdots, A_k:M),\\
&\quad M:[A_1, A_2, \cdots, A_k] = (M:A_1, \cdots, M:A_k).
\end{aligned}$$

---

[3] We use $(A, B, \cdots)$, $[A, B, \cdots]$ both for the union and join of ideals $A, B, \cdots$ or the greatest common divisor and least common multiple of integers $A, B, \cdots$.

[4] We use the notation of van der Waerden [4], chapter XII save that roman capitals are used for ideals instead of gothic capitals.

[5] A detailed analysis of the properties of residuation, is given in Ward [6], Dilworth [7].

4°. The formulas of section **3°** give the fundamental relations

$$(4.1)\quad P_a = N : L_a = \frac{N}{(N, L_a)} = \frac{[N, L_a]}{L_a},$$

$P_a = N : (N : P_a)$, $L_a = N : P_a$ if and only if L_a divides N.

We define our new operations over the group $\mathfrak{G}$ as follows. We write

$$\delta = (\xi, \eta) \text{ if } L_\delta = (L_\xi, L_\eta),$$
$$\mu = [\xi, \eta] \text{ if } L_\mu \equiv [L_\xi, L_\eta] \pmod N.$$

It is clear that the group $\mathfrak{G}$ forms an arithmetic structure [6] with respect to the operations of union and cross-cut thus defined which is simply isomorphic with the structure of the ring K_N.

The third operation over $\mathfrak{G}$ which we shall consider is a multiplication simply isomorphic with multiplication in K_N: we write

$$\pi = \xi \cdot \eta \text{ if } L_\pi \equiv L_\xi L_\eta \pmod N.$$

If we call two elements of $\mathfrak{G}$ equivalent if and only if each divides the other, then equivalent elements have the same period and conversely.

The periods of δ, μ and π obey the following simple rules which are easy consequences of (4.1) and the formulas of section **3°**:

$$P_{(\xi,\eta)} = [P_\xi, P_\eta], \quad P_{(\xi_1, \xi_2, \ldots, \xi_k)} = [P_{\xi_1}, P_{\xi_2}, \cdots, P_{\xi_k}],$$
$$P_{[\xi,\eta]} = (P_\xi, P_\eta), \quad P_{[\xi_1, \xi_2, \ldots, \xi_k]} = (P_{\xi_1}, P_{\xi_2}, \cdots, P_{\xi_k}),$$
$$P_{\xi \cdot \eta} = \{P_\xi P_\eta\} : N, \quad P_{\xi_1 \cdot \xi_2 \cdots \xi_k} = \{P_{\xi_1} P_{\xi_2} \cdots P_{\xi_k}\} : N^{k-1}.$$

Thus for each operation the period is readily calculated from the period of its constituents. It is possible to define a residual $\xi : \eta$ in $\mathfrak{G}$ by $L_{\xi:\eta} = L_\xi : L_\eta$, but its period is not calculable in terms of the periods of ξ and η alone.

5°. If we choose a different primitive element γ' in place of γ defining a new operator L'_ξ over $\mathfrak{G}$ we have

$$L'_a \equiv L'_\gamma L_a \pmod N, \qquad L_a \equiv L_{\gamma'} L'_a \pmod N$$

where $L'_\gamma L_{\gamma'} \equiv 1 \pmod N$, so that both L'_γ and $L_{\gamma'}$ are prime to N. It readily follows that the operations (ξ, η), $[\xi, \eta]$ are independent of the particular base γ chosen to define them. The situation for the product is different. We

[6] Or distributive lattice. See Ore [8] or Ward [6] for detailed definition.

find that $(\alpha \cdot \beta)' = \gamma' \cdot (\alpha \cdot \beta)$. On the other hand $P'_\alpha = P_\alpha$. The formulas (4. 1) are thus unchanged. For example:

$$P'_{(\xi \,.\, \eta)'} = P_{\gamma' \,.\, (\xi \,.\, \eta)} = \{P_{\gamma'} P_{\xi \,.\, \eta}\} : N = \{N \,.\, P_{\xi\eta}\} : N = P_{\xi\eta}.$$

6°. Suppose now that the group $\mathfrak{G}$ is the direct sum of κ cyclic groups $\mathfrak{G}^{(1)}, \cdots, \mathfrak{G}^{(k)}$ of orders $N^{(1)}, \cdots, N^{(k)}$:

$$\mathfrak{G} = \mathfrak{G}^{(1)} + \mathfrak{G}^{(2)} + \cdots + \mathfrak{G}^{(i)} + \cdots + \mathfrak{G}^{(k)}, \tag{6. 1}$$

so that the typical element α of $\mathfrak{G}$ is of the form

$$\alpha = \alpha^{(1)} + \alpha^{(2)} + \cdots + \alpha^{(i)} + \cdots + \alpha^{(k)}.$$

We select in each group $\mathfrak{G}^{(i)}$ a primitive element $\gamma^{(i)}$ and define operators $L_\alpha{}^{(i)}$ as in section **2°** by

$$L_\alpha{}^{(i)} = a^{(i)}, \qquad \alpha^{(i)} = a^{(i)} \gamma^{(i)}, \qquad (i = 1, 2, \cdots, k).$$

We then associate with the element α the vector $\mathfrak{L}_\alpha$ whose i-th component is $L_\alpha{}^{(i)}$. The operations (α, β), $[\alpha, \beta]$ $\alpha \cdot \beta$ over $\mathfrak{G}$ of union, cross-cut and product are defined by the vectors $\mathfrak{L}_{(\alpha,\beta)}$, $\mathfrak{L}_{[\alpha,\beta]}$, $\mathfrak{L}_{\alpha \,.\, \beta}$ with components $(L_\alpha{}^{(i)}, L_\beta{}^{(i)})$, $[L_\alpha{}^{(i)}, L_\beta{}^{(i)}]$, $L_\alpha{}^{(i)} L_\beta{}^{(i)}$ respectively where the components are taken modulo $N^{(i)}$ in the associated rings $K_{N^{(i)}}$.

With an obvious extension of notation, we write

$$\mathfrak{L}_{(\alpha,\beta)} = (\mathfrak{L}_\alpha, \mathfrak{L}_\beta), \qquad \mathfrak{L}_{[\alpha,\beta]} = [\mathfrak{L}_\alpha, \mathfrak{L}_\beta]$$
$$\mathfrak{L}_{\alpha \,.\, \beta} = \mathfrak{L}_\alpha \cdot \mathfrak{L}_\beta.$$

The *vectorial period* of α is defined as the vector $\mathfrak{P}_\alpha$ with components $P_\alpha{}^{(i)} = N^{(i)} : L^{(i)}$. If $\mathfrak{N}$ denotes the vector with components $N^{(1)}, N^{(2)}, \cdots, N^{(k)}$ ($\mathfrak{N}$ is simply $\mathfrak{L}_0$, where 0 is the identity element of $\mathfrak{G}$) then we write

$$\mathfrak{P}_\alpha = \mathfrak{N} : \mathfrak{L}_\alpha.$$

These definitions allow us to extend immediately the formulas of section **4°**; thus

$$\mathfrak{P}_{[\alpha,\beta]} = (\mathfrak{P}_\alpha, \mathfrak{P}_\beta), \qquad \mathfrak{P}_{(\alpha,\beta)} = [\mathfrak{P}_\alpha, \mathfrak{P}_\beta], \qquad \mathfrak{P}_{\alpha \,.\, \beta} = \mathfrak{P}_\alpha \cdot \mathfrak{P}_\beta : \mathfrak{N}.$$

The actual period of α (that is, the least positive integer p such that $p\alpha = 0$) is simply the least common multiple of the components of the vectorial period. Denoting it as before by P_α, we have

$$P_\alpha = [P_\alpha{}^{(1)}, P_\alpha{}^{(2)}, \cdots, P_\alpha{}^{(k)}].$$

We cannot calculate the scalar period of $\alpha \cdot \beta$ or $[\alpha, \beta]$ directly in terms of the scalar periods of α and β. But for the union (α, β) we have the elegant formula

$$P_{(\alpha,\beta)} = [P_\alpha, P_\beta].$$

These considerations apply to any finite Abelian group since every such group may be represented as a direct sum of cyclic groups. If we assume as is always possible that the order of each summand is a power of a prime, then the number of summands $\mathfrak{G}^{(i)}$ is uniquely specified and also the orders $N^{(i)}$.

To remove in part the ambiguity in the definition of the components of $\mathfrak{L}_\xi$ and $\mathfrak{P}_\xi$ due to the fact that the order of the groups $\mathfrak{G}^{(i)}$ in (6. 1) is unspecified, we agree to arrange the prime power orders $N^{(i)}$ first in the natural order of the primes, and then arrange the powers of each prime in order of magnitude. The remaining ambiguity in the order of the components due to adjoining isomorphic groups in the decomposition (6. 1) appears to be inherent, as the set of all vector functions $\mathfrak{P}_\xi$ over $\mathfrak{G}$ can be regarded as a basis for a representation of the group of automorphisms of $\mathfrak{G}$, each function being corollated with the sub-group of automorphisms leaving its components unchanged in order, but changing possibly the basis elements $\gamma^{(i)}$ in terms of which the components $L_\xi{}^{(i)}$ are specified. We have already seen in section **5°** that the components of $\mathfrak{P}_\xi$ are unaffected by such changes of base. The remaining automorphisms of $\mathfrak{G}$ will permute isomorphic groups in (6. 1) thus inducing a permutation of the vector functions $\mathfrak{P}_\xi$. In any event the scalar period function P_ξ remains unaffected.

Since any finite field excluding its zero element is a cyclic group with respect to multiplication, the calculus we have developed in section **4°** carries over entire to the periods of elements in any such field. The vectorial calculus of the present section similarly applies to the periods of the units in any finite commutative ring.

7°. The operations which we have defined over the finite group have analogues in common arithmetic. For if

$$A = \prod_1^\infty P_n{}^{a_n}, \qquad B = \prod_1^\infty P_n{}^{b_n}$$

are the decompositions of the positive integers A and B into their prime factors, where $P_1, P_2, P_3, \cdots$ denote the primes $2, 3, 5, \cdots$ in their natural order and only a finite number of the exponents a_n, b_n are not zero, we may define a "union," "cross-cut" and "product" of A and B "of the second kind" by

$$(A, B) = \prod_1^\infty P_n^{(a_n, b_n)} \qquad [A, B] = \prod_1^\infty P_n^{[a_n, b_n]}$$

$$A \cdot B = \prod_1^\infty P_n^{a_n, b_n}.$$

The product of the second kind is distributive with respect to the ordinary product $A \times B$ or "product of the first kind":

$$A \cdot (B \times C) = (A \cdot B) \times (A \cdot C).$$

The analogy with our treatment of groups becomes evident if we think of A as specified by the vector with components $a_1, a_2, \cdots$.

Indeed our product is the arithmetical analogue of the "multiplication of the second order" considered by De Morgan [9] and others [10] in the hierarchy of operations

$$A + B,\ A \times B = \exp(\log A + \log B),\ A \cdot B = \exp(e^{\log\log A + \log\log B}) = A^{\log B}, \cdots.$$

PASADENA, CALIFORNIA.

REFERENCES.

1. R. Dedekind, Supplement XI of Dirichlet's *Vorlesungen*, 4th edition (1894), § 170; *Collected Works*, vol. III, p. 71.
2. E. Lasker, *Mathematische Annalen*, vol. 60 (1905), pp. 20-116.
3. F. S. Macaulay, *The Algebraic Theory of Modular Systems*, Cambridge (1916).
4. B. L. van der Waerden, *Moderne Algebra*, vol. 2 (Berlin, 1931).
5. M. Ward, *Transactions of the American Mathematical Society*, vol. 35 (1933), pp. 254-260.
6. M. Ward, Paper submitted to the *Duke Journal*.
7. R. P. Dilworth, Paper submitted to the *Transactions of the American Mathematical Society*.
8. O. Ore, *Annals of Mathematics*, vol. 36 (1935), pp. 406-437.
9. A. De Morgan, *Trigonometry and Double Algebra*, p. 166.
10. R. E. Moritz, *Tohôku Journal*, vol. 21 (1921), pp. 51-64.

RATIONAL CURVES OF ORDER $n+2$ INVARIANT UNDER DIHEDRAL COLLINEATION GROUPS OF ORDER $2n$.*

By R. M. WINGER.

---

1. Introduction. The problem of self-projective curves, i. e. curves which admit linear transformations into themselves, may be said to go back to Euler and Steiner who respectively investigated curves with a diameter and a center. Projectively considered, it is a matter of indifference whether one asks that a curve have a center or a diameter. For a center of symmetry is merely the center of a harmonic homology, or reflexion, whose axis is the line at infinity; while a diameter is an axis of reflexion whose center is a point at infinity. The diameter becomes an axis of symmetry when the center of the reflexion is in the direction of the perpendicular. An algebraic curve may have several axes of symmetry although it can have but a single center. If there are two axes, the product of the corresponding reflexions will be a rotation of finite period, say p, and the reflexions will generate a dihedral group of order $2p$ which leaves the curve invariant.[1]

Multiple axial symmetry of an algebraic curve is thus intimately connected with dihedral groups. Indeed the whole question of the symmetry of curves or other geometrical figures is but a metrical aspect of collineation groups.[2]

The problem of self-projective curves, consciously formulated, has been attacked by numerous writers, including Klein and Lie, S. Kantor, Wiman, Ciani, Snyder and the Author.

The maximum axial symmetry of an algebraic curve of order m is m-fold, the axes radiating from a point at equal angles.[3] A rational m-ic with $m-1$ axes of symmetry must be of even order and a trochoid. There is in fact a one-parameter family of such curves. There is also a one-parameter family of trochoids possessing $m-2$ axes of symmetry when m is even. These two

---

* Read before the American Mathematical Society, September 3, 1936. Received by the Editors May 24, 1937.

[1] Winger, *American Mathematical Monthly*, vol. 37 (1930), p. 5. Also Loria, *Spezielle Algebraische und Transzendente Ebene Kurven*, Second Edition, vol. I. Chapter 10.

[2] Cf. Jaeger, *Lectures on the Principle of Symmetry*, Amsterdam, 1920, Chapter 3.

[3] Winger, *American Mathematical Monthly, loc. cit.*, also *American Journal of Mathematics*, vol. 36 (1914), p. 66 where the equations are obtained for rational curves.

systems have been discussed elsewhere.[4] For all values of m $(m > 3)$, even or odd however there is a one-parameter family of non-trochoidal rational curves with $m - 2$ axes of symmetry. These form the subject of the present study. We shall find it convenient to take $m = n + 2$. Projectively stated then, the problem is the consideration of those rational curves of order $n + 2$ (other than projective trochoids) which are invariant under dihedral collineation groups of order $2n$. The curves divide broadly into two classes, according as n is odd or even, corresponding to the two main species of dihedral groups. The family includes as special cases many remarkable individuals among which may be mentioned the self-dual curves of Wear,[5] which are autopolar with respect to the maximum number of conics; certain of the polar tangent curves considered by Stratton.[6] A metric version of Wear's curves are the self-dual rational curves of maximum symmetry studied by Duncan (for n odd).[7]

As usual there are two groups involved, the binary group on the parameter and the ternary group on the points of the curve. The binary group is generated by the collineation

$$s: \quad t' = \epsilon t, \qquad \epsilon^n = 1,$$

and the involution

$$r: \quad t' = 1/t.$$

The first generates an invariant cyclic subgroup g_n whose elements are $t' = \epsilon^i t$, $i = 1, 2, \cdots n$. The product of these by r yields the n involutions $t' = \epsilon^i / t$.

The ternary group is generated by the cyclic substitution

$$S: \quad x'_1 = \epsilon^2 x_1, \; x'_2 = x_2, \; x'_3 = \epsilon x_3$$

or

$$x'_1 = \epsilon x_1, \; x'_2 = \epsilon^{-1} x_2, \; x'_3 = x_3$$

of period n, which generates a ternary cyclic subgroup G_n, and the reflexion

$$R: \quad x'_1 = x_2, \; x'_2 = x_1, \; x'_3 = x_3.$$

The elements of the ternary group comprise the elements of the cyclic

$$G_n: \quad x'_1 = \epsilon^i x_1, \; x'_2 = \epsilon^{-i} x_2, \; x'_3 = x_3, \qquad (i = 1, 2, \cdots n),$$

---

[4] *American Mathematical Monthly*, vol. 39 (1932), p. 578.
[5] *American Journal of Mathematics*, vol. 51 (1929), p. 482.
[6] *American Mathematical Monthly*, vol. 43 (1936), p. 398.
[7] *Bulletin of the American Mathematical Society*, vol. 40 (1934), p. 344.

together with the products of these and R which are the n reflexions:

$$x'_1 = \epsilon^i x_2,\ x'_2 = \epsilon^{-i} x_1,\ x'_3 = x_3.$$

The axes of these reflexions are $x_1 - \epsilon^i x_2 = 0$, while their corresponding centers are $(\epsilon^i, -1, 0)$. The centers thus lie on the line $x_3 = 0$ and the axes meet at $u_3 = 0$. The equations of the axes as a whole are $x_1^n - x_2^n = 0$.

There is a fundamental difference according as n is odd or even. When n is odd the axes are all conjugate (equivalent), and no center lies on an axis. But when n is even, say $n = 2k$, they divide into two complementary sets $A+$, $A-$:

$$A+: \quad x_1^k + x_2^k = 0, \qquad A-: \quad x_1^k - x_2^k = 0,$$

the corresponding centers of which we shall denote respectively by $C+$ and $C-$. Now the axes (and centers) of either set are conjugate among themselves only. Further the axes and centers are incident in pairs: Each center $C\pm$ lies on an axis $A\mp$ of the complementary set when k is odd but on an axis $A\pm$ of its own set when k is even. Again when n is even the binary group g_n contains an extra involution, $t' = -t$, and the ternary G_n an extra reflexion with center u_3 and axis x_3. The special sets of conjugate points of the binary group are the pair of parameters $0, \infty$ and the two sets of n,

$$t^n + 1 = 0,\ t^n - 1 = 0.$$

I. The General Curve of the Pencil.

2. Consider now a rational curve of order $n+2$,

$$x_i = f_i(t), \qquad (i = 1, 2, 3),$$

where f_i are polynomials of order $n+2$, and suppose that when t is transformed into t', x (parameter t) is carried into x' (parameter t'). Then apply to f_i the generating transformations of the binary group and ask that x_i be transformed according to the generators of the ternary group. We find thus the canonical equations of the most general rational curve of order $n+2$ invariant under our group:[8]

$$x_1 = t^{n+2} + at^2,\ x_2 = at^n + 1,\ x_3 = t^{n+1} + t, \tag{1}$$

where a is an arbitrary constant. The line equations are

[8] When n is even and $n/2+1$ odd, there is a second variety in which $x_3 = t^{n/2+1}$. This is the trochoidal case already discussed, *American Mathematical Monthly*, vol. 39 (1932), p. 584.

$$
\begin{aligned}
u_1 &= at^{2n} + [n+1-(n-1)a]t^n + 1 \\
u_2 &= t^{2n+2} + [n+1-(n-1)a]t^{n+2} + at^2 \\
u_3 &= -2at^{2n+1} + [a^2(n-2) - n - 2]t^{n+1} - 2at.
\end{aligned} \tag{2}
$$

We shall first consider the properties of the curves for general parameter a, distinguishing the two cases n odd, even. While the curves are highly restricted, the treatment is quite general. We deal in a sense with a doubly infinite system of curves, since there is a one-parameter family for each value of n, which in turn may be any positive integer.[9]

The flex form is seen to factor into

$$
\begin{aligned}
&(t^n+1)\{2at^{2n} \\
&\quad + [(n-1)(n-2)a^2 - 2n^2a + (n+1)(n+2)]t^n + 2a\}.
\end{aligned} \tag{3}
$$

An immediate observation is that the first factor represents one of the special sets of conjugate points (parameters) of the binary group and it is obvious from (1) that they are cut out by x_3, hence

One third of the flexes, whose parameters comprise one of the special sets of conjugate points of the binary group, lie on the fixed line of the ternary group. The other intersections 0, ∞ of x_3 constitute the special set of two conjugate points. The remaining flexes form a general conjugate set under the binary group, since obviously the other special set $t^n - 1$ is not a factor of the flex form.

3. The general curve, n odd. The double points of the involution $t' = 1/t$ are ± 1. By referring to equations (1) we see that the parameter 1 is cut out by the axis $x_1 - x_2 = 0$. The parameter -1 names a flex whose coördinates are $(1, -1, 0)$, a center of reflexion. Further the tangent at the point $t = 1$ is, from (2), $x_1 + x_2 - (a+1)x_3 = 0$ and thus goes through the center. Since the centers as well as the axes are conjugate we may say

The centers of reflexion are flexes and thus lie on the curve. From each runs one simple tangent to the curve whose contact is on the corresponding axis. These contacts as a whole, namely $t^n - 1 = 0$ comprise the other special set of parameters under the binary group.

We have noted one point on each axis. The others must be fixed under a reflexion, while their parameters are paired under an involution. This

[9] When $n = 1$ we get a family of cubics which are all projectively equivalent except the cuspidal curve, $a = -3$.

requires nodes or multiple points of higher order. Anything but nodes however would demand more double points than are available. Hence

On each axis lie $(n+1)/2$ *nodes, which accounts for all of them.*

The nodes fall into $(n+1)/2$ special conjugate sets under the ternary group, the parameters of each set however forming general conjugate sets of the binary group. We are now in possession of all special sets of conjugate points of the ternary group which lie on the curve: Besides the nodes they are the three sets whose parameters are the special sets of the binary group.

Again, all tangents from a center of reflexion are fixed under the reflexion but their contacts are subject to the corresponding involution. This means in the general case that all tangents from a center, except the flex tangent of the center itself and the simple tangent already noted, must be bitangents. We may say then, the curve being of class $2n+2$,

$n-1$ *bitangents meet at each center of reflexion, accounting for half the total.*

4. The general curve, n even. When n is even the changes we have to record are due mainly to the division of the axes (and centers) into two sets which are not conjugate to each other, and to the presence of an additional reflexion, whose center is u_3 and whose axis is x_3. The double points of the associated involution, $t' = -t$, are $0, \infty$, which lie on the axis and name contacts of tangents from the center. The other intersections of this axis will be nodes but, since they are also flexes, biflecnodes are formed. Hence

There are $n/2$ *biflecnodes on the axis* $x_3 = 0$.

Now setting $n = 2k$, we have to distinguish two cases according as k is odd or even.

k odd: The axis $x_1 - x_2 = 0$ cuts out the pairs of parameters $t^2 - 1 = 0$ and $t^2 + 1 = 0$, which are double points of two involutions. The first pair name contacts of tangents from the center of the reflexion, while the others are parameters of a biflecnode on x_3. The remaining intersections of the axis are ordinary nodes. The biflecnode is a center of reflexion whose axis is $x_1 + x_2 = 0$. All intersections of this axis are nodes, for the double points of the associated involution fall at the center. And all the tangents from a biflecnode, except the nodal tangents themselves, are double lines. Since these results are typical, we have

Each of the axes $A-$ cuts out one biflecnode, $(n-2)/2$ ordinary nodes, and two contacts of simple tangents from its associated center; whereas each axis $A+$ cuts out $(n+2)/2$ ordinary double points. From each biflecnode run $n-2$ double tangents, but from each of the other centers, including u_3, run two simple and n double tangents.

This accounts for all of the double points and n^2 double lines, leaving $n(n-2)$ double lines to be distributed in general sets.

k even: Each of the axes $A-$ cuts out $n/2$ ordinary double points and two contacts of tangents from the corresponding center, while each of the complementary sets of axes $A+$ cuts out one biflecnode and $n/2$ ordinary double points, accounting for all of the double points.

Each biflecnode is a center $C+$ and thus lies on an axis conjugate to its own. Through each biflecnode pass $n-2$ bitangents and through each of the other centers, including u_3, pass two simple and n double tangents, accounting for a total of n^2 bitangents as before.

5. The pencil of invariant conics. Whether n is odd or even, there is a pencil of conics which are individually invariant under the group. The equation of the pencil in points and in lines may be written

$$x_1x_2 = \lambda x_3^2 \text{ and } 4\lambda u_1u_2 = u_3^2. \tag{4}$$

All proper members of the pencil touch the curve at the points $t = 0, \infty$ and cut out besides $2n$ points which comprise conjugate sets, special or general, under both groups. Likewise the $4n$ common lines of each conic and the curve form conjugate sets of both groups. Indeed the common points and lines of the curve and degenerate members of both pencils belong in conjugate sets of the groups. Conversely, all conjugate sets of points and lines of both groups, special as well as general, are common points and lines of the conics and the curve. Among the conics of particular interest are:

	One conic	Value of λ
(5)	with contacts at $t^n = 1$,	$\dfrac{(a+1)^2}{4}$
(6)	on $2n$ flexes,	$\dfrac{na[(n-3)a-(n+3)]}{(n-1)(n-2)a-(n+1)(n+2)}$
(7)	on flex lines[10] $t^n + 1 = 0$,	$\dfrac{[(n-2)a+n+2]^2}{4n^2}$

[10] Biflecnodal tangents when n is even.

(8) on the other $2n$ flex lines, $\dfrac{na[(n-2)a-(n+2)]}{(n-1)^2a-(n+1)^2}$

(9) on contacts of tangents [11] from u_3, $\dfrac{na(a+1)}{(n-2)a+n+2}$.

We note that the value of λ in (5) is independent of n. We observe also that two or more of the conics may coincide or one or more of the conics may degenerate ($\lambda = 0, \infty$) for particular values of a, i. e. for certain special curves. On the other hand, the coincidence of any two of the conics or the degeneration of any one implies a specialization of the curve.

II. Special Cases.

6. By assigning particular values to a we obtain an infinite system of projectively distinct curves. Among the most interesting special curves are those with singularities arising from the coincidence of the flexes. These are found by equating to zero the discriminant of the second factor of the flex form (3).[12] We shall consider them briefly. The discriminant of the second factor of the flex form, considered as a quadratic, is

$$(a-1)[(n-1)a-(n+1)][(n-2)a-(n+2)] \times [(n-1)(n-2)a-(n+1)(n+2)].$$

The complete discriminant is a power of this, together with two other factors corresponding to $a = 0, \infty$. For obviously these values imply multiple roots of the flex equation. When $a = 1, \infty$ the curve degenerates for then the three line sections $x_i = 0$ have a common factor. Ordinarily the flexes in question form a general set under the binary group. But when the discriminant vanishes the flex parameters coincide at least in pairs and hence must reduce to special sets. We find readily the following consequences of a vanishing discriminant:

When	The curve has
$a = (n+1)/(n-1)$,	undulations at $t^n - 1 = 0$
$a = (n+2)/(n-2)$,	cusps at $t^n - 1 = 0$
$a = (n+1)(n+2)/(n-1)(n-2)$,	5-point contact flexes at $t^n + 1 = 0$
$a = 0$,	n-point contact tangents at $t = 0, \infty$.

In the first three cases the flexes involved coincide in pairs at the respective

[11] Bitangents when n is even.

[12] The flexes represented by the first factor obviously cannot coincide.

singularities, but when $a = 0$, they coincide n at a time. We shall now discuss the individual curves.

$$a = (n+1)/(n-1).$$

7. The undulations lie on the axes of reflexion, one on each when n is odd and two on each axis $A-$ when n is even. The undulation tangents pass through the centers of reflexion, one through each when n is odd and two through each $C-$ center when n is even. The other properties of the curve noted for the general case must be modified to conform to the fact that an undulation arises from the coincidence of three lines through a center of reflexion, one bitangent, one line joining two flexes, and one simple tangent.

$$a = (n+2)/(n-2).$$

8. This is the self-dual case treated by Wear (l. c.). The cusps reduce the class of the curve to $n+2$ and the number of double lines to $n(n-1)/2$. When n is odd, there is one cusp on each axis, which is the cusp tangent. Also $(n-1)/2$ double lines meet at each flex (center of reflexion). When n is even however the cusps fall in pairs on the axes $A-$, which are now double-cusp tangents and count as double lines. Further when n is even and k is odd, each double-cusp tangent cuts out one of the biflecnodes and $(n-6)/2$ ordinary nodes; while the other axes $A+$ cut out $(n+2)/2$ ordinary nodes each. Further $(n-4)/2$ double lines, including one double cusp tangent, meet at each biflecnode and $(n+2)/2$ double lines meet at each of the other centers $C-$. When k is even, each double cusp tangent cuts out $(n-4)/2$ ordinary nodes, whereas the other axes $A+$ cut out $n/2$ ordinary nodes and one biflecnode each. Hence in this case $(n+2)/2$ double lines, including one double cusp tangent, meet at each center $C-$, while $(n-4)/2$ double lines meet at each center $C+$, which is now a biflecnode. When n is even, the tangents from the additional center u_3 consist of two simple tangents and the k double cusp tangents, which have already been enumerated. All the singularities of the self-dual curves have now been accounted for.

The line equations of the curve, after factoring out the cusp form, reduce to

$$\begin{aligned} u_1 &= (n+2)t^n - (n-2) \\ u_2 &= (n-2)t^{n+2} - (n+2)t^2 \\ u_3 &= -2(n+2)t(t^n-1). \end{aligned} \tag{10}$$

The binary transformation $t' = \eta/t$, where $\eta^n = -1$ induces the polarity

$$\Pi: \quad u_1 = x_1,\ u_2 = x_2,\ u_3 = -2a\eta x_3, \qquad a = (n+2)/(n-2)$$

which interchanges the dual singularities of the curve. This combined with the G_{2n} yields the $2n$ correlations:

$$(11)\qquad \begin{aligned} &u_1 = \epsilon^i x_1,\ u_2 = \epsilon^{1-i} x_2,\ u_3 = -2a\eta x_3, \\ &u_1 = \epsilon^i x_2,\ u_2 = \epsilon^{1-i} x_1,\ u_3 = -2a\eta x_3, \end{aligned} \qquad (i = 1, 2, \cdots n).$$

The first set of these are polarities in all cases. Moreover when n is odd, say $n = 2p-1$, we get a polarity in the second set, namely when $i = p$. We may summarize thus:

When $a = (n+2)/(n-2)$, the curve admits a G_{4n}, comprising $2n$ collineations and $2n$ correlations. Of the correlations n are polarities when n is even while $n+1$ are polarities when n is odd. The curve is thus autopolar with respect to n or $n+1$ conics, according as n is even or odd—the maximum number for a rational curve of order $n+2$.

$$a = (n+1)(n+2)/(n-1)(n-2).$$

9. All of the flexes of this curve unite to form higher flexes with 5-point contact tangents at $t^n + 1 = 0$, distinct when n is odd but joined at special biflecnodes when n is even. These singularities absorb a total of $3n$ double lines and reduce the number of tangents that can be drawn from each of them by 5 or 10, according as n is odd or even. The properties of the general curve will be modified in consequence. Hence, when n is odd the number of double lines that meet at each flex in question is $n-2$, leaving $n(n-3)$, which belong in general sets. When n is even, k odd or even, $n-4$ double lines meet at each biflecnode and n at each of the other $k+1$ centers, $n(n-4)$ remaining.

$$a = 0.$$

10. The equations reduce to

$$(12)\qquad x_1 = t^{n+2},\ x_2 = 1,\ x_3 = t^{n+1} + t.$$

There are hyperosculation points at $t = 0, \infty$, whose tangents are $x_1 x_2 = 0$ and which absorb together $2n$ flexes and $n(n-1)$, i. e. half of the double lines. The other intersections of x_3 are simple flexes or ordinary biflecnodes according as n is odd or even. Thus all the flexes lie on x_3. All tangents from u_3 coincide with $x_1 x_2 = 0$. The pair of lines $x_1 \pm x_2 = 0$ are cyclic cutting out the points $t^{n+2} \pm 1 = 0$.

The dual of the line curve may be written, see equations (2),

$$(13) \qquad \begin{aligned} x_1 &= t^{2n+2} + (n+1)t^{n+2}, \\ x_2 &= (n+1)t^n + 1, \\ x_3 &= -(n+2)t^{n+1}. \end{aligned}$$

This is a projective trochoid having an $(n+1)$-fold point with coincident parameters at each of the points $t = 0, \infty$. Each of these points, which is the dual of a tangent with $(n+2)$-point contact, is thus equivalent to n cusps and $n(n-1)/2$ nodes.

Numerous other noteworthy curves occur in the family, among the most interesting of which are those arising when some of the invariant conics of § 5 degenerate or coincide. We get degenerate conics (and special curves) when $\lambda = 0$ or ∞. These cases we shall now consider.

$$a = -1.$$

11. This curve has an n-fold point at u_3 with parameters $t^n - 1 = 0$. When n is odd the multiple point has distinct tangents and absorbs $n(n-1)/2$ double points, leaving n, one of which lies on each axis. Moreover the tangent at each branch is on one of the centers of reflexion. When n is even however, the tangents touch in pairs, forming k tac-nodes within the multiple point which absorb $n/2$ extra double points and a like number of double lines. These with the usual $n/2$ biflecnodes account for all of the nodes. The tac-nodal tangents are the axes $A+$ or $A-$ according as k is odd or even and are thus on the centers $C-$ in both cases. Each counts as a bitangent and replaces the two simple tangents from a center $C-$ of the general case. Hence, k odd or even, all of the tangents from each center $C-$ are double lines. From each center $C+$ (a biflecnode) run $n-2$ double lines. We have now accounted for n^2 bitangents, including those absorbed by the tac-nodes. The center u_3 yields no new ones since all of the double lines from there are the tac-nodal tangents. Conics (5) and (9) in lines both degenerate to $u_3^2 = 0$, i. e. the multiple point repeated.

$$a = (n+3)/(n-3).$$

12. Conic (6) reduces to $x_3^2 = 0$ when

$$a = (n+1)(n+2)/(n-1)(n-2)$$

for then all the flexes combine to form higher flexes (supra). This conic also degenerates when $a = 0$ or $(n+3)/(n-3)$, reducing to $x_1x_2 = 0$. The case $a = 0$ has already been noticed. For the other value of a the flex form factors into

$$[(n-3)t^n + n + 3][(n+3)t^n + n - 3]$$

and we have at once:

If $n > 3$ and $a = (n+3)/(n-3)$, the $3n$ flexes are cut out by the sides of the invariant triangle, one third lying on each side. When n is odd all the flexes are distinct but when n is even those on $x_3 = 0$ form biflecnodes as usual.[13]

$$a = -(n+2)/(n-2).$$

13. Conic (7) in lines degenerates to $u_3^2 = 0$ for this value of a which is the negative of that for the self-dual case. The tangents from u_3 by (2)) are now given by $t(t^n+1)^2 = 0$, hence in this case

The tangents from the fixed point of the ternary group, except the simple tangents $0, \infty$, are all flex tangents with their contacts on the fixed line. If n is odd they are all simple flex lines, but if n is even each is a tangent at a special biflecnode whose tangents coincide.

Each such biflecnode or oscnode counts for three ordinary double points, i. e. contains two latent bitangents. Since an ordinary double flex tangent is equivalent to two flex and four double tangents, these oscnodal tangents are equivalent to two flex and six double tangents. Thus the tangents from u_3 account for $3n$ bitangents. The oscnodes are centers of reflexion $C+$, lying on the axes $A-$ or $A+$ according as k is odd or even. From each can be drawn to the curve, besides its own tangent, $2n-4$ tangents which comprise $n-2$ double lines; while from each of the centers $C-$ can be drawn 2 simple and n double tangents. We have now accounted for $n(n+2)$ bitangents, $n(n-4)$ remaining.

$$a = (n+1)^2/(n-1)^2.$$

14. For the self-dual curves, as well as when $a = 0$, the line conic (8) reduces to $u_3^2 = 0$, i. e. the point (repeated) in which the cusp or the hyperosculation tangents respectively meet. When a takes the value above the conic becomes $u_1 u_2 = 0$, whose common lines with the curve, except $0, \infty$, are given by

$$[(n+1)t^n - n + 1]^2[(n-1)t^n - n - 1]^2 \equiv (\text{second factor of flex form})^2.$$

[13] It is a rather common experience to encounter a dihedral curve with numerous flexes on a line at the centers of reflexion. The trinomial curves $x^n + y^n + z^n = 0$, n a positive integer, meet the sides of the reference triangle, which are axes of homology, in hyperosculation points which absorb all of the flexes. But this is the first non-trivial instance I know of a rational curve whose full complement of simple flexes lie on the sides of a proper triangle.

Hence *n flex tangents meet at each of the points* $u_1 = 0$, $u_2 = 0$. Or,

When $a = (n+1)^2/(n-1)^2$, *one third of the inflexional lines meet at each of the points cut out by the line on which lie the contacts of the remaining third of inflexional lines.*

$$a = -(n+1)/(n-1).$$

15. *If n is odd and* $a = -(n+1)/(n-1)$, *the tangents from the centers of reflexion (flexes) touch the curve again on the axes, at the points* $t^n - 1 = 0$.

This fact, first observed for the quintic [14] is easily verified for the general case. These line singularities are triple lines, each equivalent to one flex tangent and two bitangents, accounting thus for a total of $2n$ bitangents. From each center can be drawn $2n-2$ other tangents which must unite to form $n-1$ bitangents. This accounts for $n(n+1)$, so that $n(n-3)$ are left.

What is the meaning when n is even? The biflecnodal tangents cannot touch the curve at one of the points $t^n - 1 = 0$ for these points do not lie on the axes of which the biflecnodes are centers. Now we saw that there is a conic (5) which touches the curve at the points $t^n - 1 = 0$; and when n is odd the flex lines $t^n + 1 = 0$ touch the curve at the same points. Hence the conic (7) on these flex lines must be identical with conic (5). In fact the condition that these conics be the same for a proper curve is just that $a = -(n+1)/(n-1)$ whether n be odd or even, the equations reducing to

$$4u_1u_2 = (n-1)^2u_3^2, \text{ or } (n-1)^2x_1x_2 = x_3^2.$$

For the even case this conic is on the biflecnodal lines and we may summarize as follows:

If $a = -(n+1)/(n-1)$, *the conic on the flex lines* $t^n + 1 = 0$ *touches the curve at the points* $t^n - 1 = 0$ (*as well as at* $0, \infty$). *When n is odd the flex tangents touch curve and conic at the same points and thus count for four common lines each. When n is even the flex lines in question are biflecnodal tangents, which however do not touch the curve elsewhere. In either case all common lines of curve and conic are comprised in the flex* (*or biflecnodal*) *tangents and the tangents at the contacts.*

---

[14] Winger, "Self-projective rational curves of the fourth and fifth orders," *American Journal of Mathematics*, vol. 36 (1914), p. 75, where the curve is taken in slightly different form.

III. Metrical Specializations.

16. All of the foregoing curves can be given a metrical setting. The most effective way to do this is to introduce circular coördinates x, $\bar{x}$ by means of the relations

$$\begin{aligned} x &= X + iY = x_1/x_3 \\ \bar{x} &= X - iY = x_2/x_3 \end{aligned} \tag{14}$$

where X, Y are ordinary Cartesian coördinates and t is to be considered as a complex number of absolute value 1.[15] The general equation of the curve now becomes [16]

$$x = (t^{n+1} + at)/(t^n + 1), \qquad t = \cos\theta + i\sin\theta \equiv \operatorname{cis}\theta. \tag{15}$$

In this form the axes of reflexion are axes of symmetry equi-spaced about a point, the Cartesian origin. When n is even this point is a center of symmetry, i. e. a center of the curve. All of the curves are circular, touching the circular rays at I and J. The invariant conics (when proper) are concentric circles. When $a = (n+2)/(n-2)$, the self-dual case, and n is odd we get the curves treated by Duncan (l. c.).

To write the parametric equations of (15) in Cartesian coördinates, we have, expressing t in its trigonometric form

$$x = [\operatorname{cis}(n+1)\theta + a\operatorname{cis}\theta]/(1 + \operatorname{cis} n\theta).$$

Changing the denominator to half angle functions, we get

$$x = [\operatorname{cis}(n+1)\theta + a\operatorname{cis}\theta]/[2\cos(n\theta/2)\operatorname{cis}(n\theta/2)],$$

whence, dividing by $\operatorname{cis}(n\theta/2)$,

$$x \equiv X + iY = \{\operatorname{cis}[(n+2)\theta/2] + a\operatorname{cis}[(2-n)\theta/2]\}/2\cos(n\theta/2).$$

Equating real and imaginary parts, we have the desired equations:

$$\begin{aligned} 2X &= \{\cos[(n+2)\theta/2] + a\cos[(n-2)\theta/2]\}/\cos(n\theta/2) \\ 2Y &= \{\sin[(n+2)\theta/2] - a\sin[(n-2)\theta/2]\}/\cos(n\theta/2). \end{aligned} \tag{16}$$

---

[15] One equation suffices for this representation, though associated with it is the conjugate $\bar{x} = (at^n + 1)/(t^{n+1} + t)$.

[16] This is really a matter of interpretation rather than a transformation of coördinates. We should have been led to the same result had we written the equations of the curve and the generators of the group in circular coördinates to begin with and then restricted the curve to admit the group.

17. $a = -1$. Another striking metrical case is the curve for which $a = -1$. If we set $a = -1$ in (16) and apply trigonometric formulas for changing sums to products, we get

$$X = -\tan(n\theta/2)\sin\theta, \qquad Y = \tan(n\theta/2)\cos\theta.$$

Rotating the axes through 90°, we have $X = -Y'$, $Y = X'$ and the equations become, dropping primes,

$$X = \tan(n\theta/2)\cos\theta, \qquad Y = \tan(n\theta/2)\sin\theta, \tag{17}$$

whence we write down at once the polar form

$$\rho = \tan(n\theta/2). \tag{18}$$

Therefore, *if* $a = -1$ *our curves are projectively equivalent to the polar tangent curves* (18).[17]

Or, reversing the last step, we may say that *the curve* $\rho = \tan(n\theta/2)$, *where* n *is an integer* (*not zero*), *is a rational curve of order* $n + 2$ *belonging to the class of curves of this paper.*

18. Again, let $a = 0$. We have at once from (16), putting $(n+2)\theta/2 = \phi$

$$\begin{aligned} 2X &= \sec[n\phi/(n+2)]\cos\phi \\ 2Y &= \sec[n\phi/(n+2)]\sin\phi \end{aligned} \tag{19}$$

which in polar form becomes

$$2\rho = \sec[n\phi/(n+2)]. \tag{20}$$

In this form the curve, which is obviously the inverse with respect to the pole of the rose curve $\rho = \cos[n\phi/(n+2)]$, has been widely studied under the name of the epi or Cotes's spiral.[18] Hence,

When $a = 0$, *our curve may be projected into the epi* (20).

UNIVERSITY OF WASHINGTON,
SEATTLE, WASHINGTON.

---

[17] The curves $\rho = a\tan m\theta + k$, where m is rational, have been studied by Stratton, *American Mathematical Monthly*, vol. 43 (1936), p. 398 ff. He discusses the symmetry but not the implied group properties of these curves.

[18] For reference see Loria, *Spezielle Algebraische und Transzendente Ebene Kurven*, I, pp. 366, 367.

CONTRIBUTION À L'ÉTUDE DES SYSTÈMES DE CHOSES NORMÉES.*

Par V. GLIVENKO.

I. Position du problème.

1. Dans un article récent,[1] j'ai étudié les systèmes de choses normées, c'est à dire les ensembles S d'éléments $a, b, \cdots$ satisfaisant aux axiomes suivants:

Axiomes des choses.

1°. L'ensemble S contient des couples d'éléments a, b liés entre eux par une relation $a \subset b$ telle que $a \subset b$ et $b \subset a$ entraîne $a = b$ et inversement, et que $a \subset b$ et $b \subset c$ entraîne $a \subset c$.

2°. À tout couple d'éléments a, b de l'ensemble S correspond un élément ab, de S, tel que $ab \subset a$, $ab \subset b$ et que $x \subset a$ et $x \subset b$ entraîne $x \subset ab$.

3°. À tout couple d'éléments a, b de l'ensemble S correspond un élément $a + b$, de S, tel que $a \subset a + b$, $b \subset a + b$ et que $a \subset y$ et $b \subset y$ entraîne $a + b \subset y$.

4°. L'ensemble S contient un élément 0 tel que, quel que soit l'élément z de S, on a $0 \subset z$.

Axiome de la norme.

À tout élément a de l'ensemble S correspond un nombre non négatif $| a |$, norme de cet élément, tel que $a \subset b$ et $a \neq b$ entraîne $| a | < | b |$, qu'on a

$$| a + b | + | ab | = | a | + | b |$$

et qu'on a $| 0 | = 0$.

Actuellement, nous ne considérons qu'un cas particulier des dits systèmes, que nous appellerons systèmes de choses *complètement* normées et qui satisfont, par définition, aux axiomes supplémentaires suivants:

Axiome supplémentaire des choses.

5°. L'ensemble S contient un élément 1 tel que, quel que soit l'élément z de S, on a $z \subset 1$.

* Received January 15, 1937.

[1] "Géometrie des systèmes de choses normées," *American Journal of Mathematics*, vol. 58 (1936), pp. 799-828.

Axiome supplémentaire de la norme.

On a $| 1 | = 1$.

Nous appelons systèmes distributifs les systèmes où, pour tous les trois choses a, b, c a lieu la loi distributive dans la forme

$$ac + bc = (a + b)c$$

ou bien dans la forme

$$(a + c)(b + c) = ab + c$$

qui est équivalente à la précédente.

2. Dans l'article cité, j'ai introduit la notion d'espace métrique presque ordonné. C'est l'espace métrique D contenant un point que nous appelons origine et qui possède les propriétés suivantes (1′ et 2′). Rappelons que, (a, b) désignant la distance de point a et de point b, on dit qu'un point c se trouve entre deux points a et b si l'on a

$$(a, c) + (c, b) = (a, b).$$

Convenons de dire maintenant qu'un point a est plus prochain qu'un point b, ou bien que b est plus lointain que a, si a se trouve entre l'origine et b. Alors:

1′. Si les points a et b, de D, sont plus prochains qu'un point x, chaque point qui se trouve entre a et b est, lui-aussi, plus prochain que x; de même, si les points a et b sont plus lointains que y, chaque point qui se trouve entre a et b est, lui-aussi, plus lointain que y.

2′. Parmi les points, de D, qui se trouvent entre les deux points donnés quelconques, il existe un qui est le plus prochain et il existe un autre qui est le plus lointain.

J'ai établi que tout système S de choses normées est un espace métrique presque ordonné où la distance de a et de b est égale à

$$| a + b | - | ab |.$$

Nous appelons espaces transitifs les espaces métriques presque ordonnés où la condition suivante est remplie:

T. Si un point c, de D, se trouve entre x et y et si tous les deux points x et y se trouvent entre a et b, le point c se trouve, lui-aussi, entre a et b.

Convenons de dire, pour abréger, que le système S est *métrisé*, lorsqu'on a pris l'expression $| a + b | - | ab |$ pour la distance de a et de b. J'ai établi que tout système métrisé distributif S, de choses normées, est un espace métrique presque ordonné transitif.

Tous les deux résultats mentionnés, concernant les relations entre les systèmes de choses normées et les espaces métriques, ont été obtenus en prenant toujours la chose 0 du système pour l'origine de l'espace correspondant. Or, il est naturel de poser le problème s'il est possible, ou non, de prendre une autre chose $u \neq 0$ pour l'origine, en sorte qu'il y restent intactes les conditions 1′ et 2′ (dans le cas général) et la condition T (dans le cas des systèmes distributifs).

La solution de ce problème étant intéressante grâce à la simplicité des énoncés définitifs, je suis décidé à la publier ici.

II. Condition de Dedekind.

3. Dans tout ce qui suit, il nous sera d'une grande importance un fait qui a été bien éclairci, entre autres, dans les travaux de M. Garrett Birkhoff.[2] C'est que *tout système S de choses normées satisfait à la condition de Dedekind.* Cette dernière peut s'énoncer comme il suit:

(D_1) Quelles que soient les trois choses a, b, c de S, où $a \subset c$, on a

$$(a + b)c = a + bc.$$

On peut lui attribuer aussi une autre forme que voici:

(D_2) Quelles que soient les trois choses a, b, c de S, on a

$$(ac + b)c = ac + bc.$$

On établit sans peine l'équivalence de (D_1) et de (D_2) en remarquant que, pour qu'on ait $a \subset c$, il faut et il suffit qu'on ait $ac = a$.

Démontrons maintenant que tout système S de choses normées satisfait à (D_2). En premier lieu, on a toujours

$$ac + bc \subset (ac + b)c.$$

Pour s'en convaincre, il suffit de remarquer qu'on a $ac \subset ac + b$ et $ac \subset c$, par suite $ac \subset (ac + b)c$, et qu'on a $bc \subset ac + b$ et $bc \subset c$, par suite $bc \subset (ac + b)c$. Il nous reste donc à établir l'égalité

$$| ac + bc | = | (ac + b)c |.$$

On s'appuie ici sur un principe général. En effet, s'il était quelque part $x \subset y$ et $x \neq y$, il serait nécessairement $| x | < | y |$. Donc, de $x \subset y$ et $| x | = | y |$ il s'ensuit toujours $x = y$.

Pour établir l'égalité $| ac + bc | = | (ac + b)c |$, il suffit d'effectuer un simple calcul, à savoir:

---

[2] "On the combination of subalgebras" et "Applications of lattice algebra," *Proceedings of the Cambridge Philosophical Society*, vol. 23 (1933), pp. 441-469, et vol. 30 (1934), pp. 115-112 respectivement.

$$\begin{aligned}|(ac+b)c| &= |ac+b|+|c|-|ac+b+c| \\ &= |ac|+|b|-|acb|+|c|-|b+c| \\ &= |ac|-|acbc|+|bc| = |ac+bc|.\end{aligned}$$

4. Dans l'article cité, j'ai démontré que, dans l'espace formé par un système métrisé S, un point c se trouve entre deux points a et b si et seulement si l'on a, dans S,

$$(E_1) \qquad ac+bc = c = (a+c)(b+c).$$

Grâce à la condition de Dedekind, cette condition (E_1) peut se présenter aussi sous une forme distincte. En effet, en se servant de la condition de Dedekind, on obtient:

$$ac+bc = (ac+b)c \quad \text{et} \quad (a+c)(b+c) = c+(a+c)b.$$

Par conséquent, les égalités (E_1) peuvent s'écrire

$$(ac+b)c = c = c+(a+c)b,$$

ou, ce qui revient au même,

$$(E_2) \qquad (a+c)b \subset c \subset ac+b.$$

C'est la forme en question.

III. La notion d'extrémité.

5. Appelons *extrémité* d'un système S chaque chose u, de S, qui possède les deux propriétés suivantes:

1) A la chose u correspond une autre chose $\tilde{u}$, de S, telle que

$$(1) \qquad u\tilde{u} = 0,$$

$$(2) \qquad u+\tilde{u} = 1.$$

2) Quelles que soient les deux choses a et b, de S, on a

$$(3) \qquad (a+b)u = au+bu,$$

$$(4) \qquad (a+b)\tilde{u} = a\tilde{u}+b\tilde{u}.$$

Il est évident que, u étant une extrémité, $\tilde{u}$ l'est aussi; nous dirons que u et $\tilde{u}$ sont les extrémités *opposées*.

Établissons tout d'abord que, u étant une extrémité d'un système métrisé S, de choses complètement normées, on a, quelles que soient les deux choses a et b, de S,

$$ab + u = (a + u)(b + u).$$

Puisqu'on a toujours $ab + u \subset (a + u)(b + u)$, pour établir l'égalité qui vient d'être écrite, il suffit de démontrer l'égalité

$$|\, ab + u \,| = |\,(a + u)(b + u)\,|.$$

A cet effet, remarquons qu'on a

$$|\,(a + u)(b + u)\,| = |\, a + u \,| + |\, b + u \,| - |\, a + b + u \,|.$$

Or,

$$|\, a + u \,| + |\, b + u \,| = |\, a \,| + |\, u \,| - |\, au \,| + |\, b \,| + |\, u \,| - |\, bu \,|$$

et, en tenant compte de (3),

$$|\, a + b + u \,| = |\, a + b \,| + |\, u \,| - |\,(a + b)u\,|$$
$$= |\, a + b \,| + |\, u \,| - |\, au + bu \,| = |\, a + b \,| + |\, u \,| - |\, au \,| - |\, bu \,| + |\, abu \,|.$$

Donc,

$$|\,(a + u)(b + u)\,| = |\, a \,| + |\, b \,| - |\, a + b \,| + |\, u \,| - |\, abu \,|.$$

Or,

$$|\, a \,| + |\, b \,| - |\, a + b \,| = |\, ab \,|$$

et

$$|\, u \,| - |\, abu \,| = |\, u \,| - |\, ab \,| - |\, u \,| + |\, ab + u \,| = |\, ab + u \,| - |\, ab \,|.$$

Donc,

$$|\,(a + u)(b + u)\,| = |\, ab + u \,|.$$

Remarque. Le fait que u est une extrémité n'y joue, de fait, aucun rôle. Quelle que soit la chose déterminée c, de S, l'égalité

$$(5) \qquad (a + b)c = ac + bc$$

a pour conséquence l'égalité

$$(6) \qquad ab + c = (a + c)(b + c).$$

Nous venons de voir que ceci a lieu dans les systèmes S de choses normées. M. O. Ore [3] a réussi à établir une proposition plus générale, á savoir que l'égalité (6) est une conséquence de (5) dans tous les systèmes où la condition de Dedekind est remplie.

6. THÉORÈME I. *La chose u étant une extrémité d'un système S de choses complètement normées, son opposée $\tilde{u}$ ne peut être définie que d'une manière univoque.*

[3] "On the foundation of abstract algebra," I, *Annals of Mathematics*, vol. 36 (1935), p. 416.

Démonstration. La chose u étant une extremité, on a, quelles que soient les deux choses a et b,

$$(a+b)u = au + bu$$

et, comme nous l'avons vu tout à l'heure,

$$ab + u = (a+u)(b+u).$$

Prenons maintenat les choses x et y qui possèdent, tous les deux, les propriétés de l'extrémité $\tilde{u}$, de sorte qu'on a, en particulier,

$$ux = 0, \qquad uy = 0,$$
$$u + x = 1, \qquad u + y = 1.$$

Il en résulte qu'on a

$$(x+y)u = xu + yu = 0, \qquad xy + u = (x+u)(y+u) = 1.$$

On en obtient, en se servant de (D_1),

$$xy = xy + (x+y)u = (xy+u)(x+y) = x + y,$$

d'où $x = y$.

THÉORÈME II. *La chose u étant une extrémité d'un système métrisé S, de choses complètement normées, la condition nécessaire et suffisante pour que, dans l'espace formé par S, un point a se trouve entre u et un autre point b, est qu'on ait*

$$ub \subset a \subset u + b.$$

Démonstration. D'après (E_2), a se trouve entre u et b si et seulement si l'on a

$$(b+a)u \subset a \subset ba + u.$$

Or, ceci équivaut à

$$bu + au \subset a \subset (b+u)(a+u),$$

ce qui equivaut, à son tour, à

$$ub \subset a \subset u + b.$$

THÉORÈME III. *Les choses u et $\tilde{u}$ étant deux extrémités opposées d'un système métrisé S, dans l'espace formé par S chaque point z se trouve entre u et $\tilde{u}$.*

Démonstration. En vertu du théorème II, pour que z se trouve entre u et $\tilde{u}$, il suffit qu'on ait

$$u\tilde{u} \subset z \subset u + \tilde{u}.$$

Or, ceci est toujours vrai, parce que $u\tilde{u} = 0$ et $u + \tilde{u} = 1$.

THÉORÈME IV. *Les choses u et $\tilde{u}$ étant deux extrémités opposées d'un système métrisé S, si, dans l'espace formé par S, un point a se trouve entre u et b, alors b se trouve entre a et $\tilde{u}$ et réciproquement.*

Démonstration. En vertu du théorème II, si a se trouve entre u et b, on a

$$ub \subset a \subset u + b.$$

Il s'ensuit que, d'une part,

$$\tilde{u}a \subset \tilde{u}(u + b) = \tilde{u}u + \tilde{u}b = \tilde{u}b \subset b$$

et que, d'autre part,

$$b \subset \tilde{u} + b = (\tilde{u} + u)(\tilde{u} + b) = \tilde{u} + ub \subset \tilde{u} + a.$$

En somme, on a

$$\tilde{u}a \subset b \subset \tilde{u} + a,$$

ce qui suffit, d'après le théorème II, pour que b se trouve entre a et $\tilde{u}$. La réciproque se démontre, naturellement, de même.

IV. Théorème fondamental.

7. Il serait très commode d'avoir des symboles spéciaux pour le plus prochain et le plus lointain des points que se trouvent entre les deux points donnés, a et b. Dans la suite, nous *employerons les symboles* $a \wedge b$ (pour le point le plus prochain) *et* $a \vee b$ (pour le point le plus lointain).

Ceci posé, le théorème fondamental de cet article s'énoncera comme il suit:

THÉORÈME V. *Dans l'espace formé par un système métrisé S de choses complètement normées, un point u peut être pris pour l'origine si et seulement si la chose u est une extrémité de S. Dans ce cas, on aura, en désignant par $\tilde{u}$ l'extrémité opposée de u:*

$$(7)\quad a \wedge b = ab + u(a + b)$$
$$= ab + ua + ub = (a + b)(u + ab) = (a + b)(u + a)(u + b),$$

$$(8)\quad a \vee b = ab + \tilde{u}(a + b)$$
$$= ab + \tilde{u}a + \tilde{u}b = (a + b)(\tilde{u} + ab) = (a + b)(\tilde{u} + a)(\tilde{u} + b).$$

Démonstration. Nous utiliserons partout le fait que, si un point c se trouve entre deux points a et b, on a nécessairement

$$ab \subset c \subset a + b$$

(ceci est immédiat, car, si c se trouve entre a et b on a, d'après (E_1),

$$ab \subset (a + c)(b + c) = c = ac + bc \subset (a + b)$$

et que, lorsque l'un des points a ou b est une extrémité, la réciproque est vraie elle-aussi (ceci résulte du théorème II).

Soit maintenant S un système métrisé, de choses complètement normées, et soit u une extrémité de ce système. Essayons de prendre u pour l'origine de l'espace formé par S. Alors, par définition, un point a sera plus prochain qu'un autre point b, ou bien, b sera plus lointain que a, si et seulement si a se trouvera entre u et b. Nous allons voir que tous les deux conditions 1′ et 2′ y seront remplies et que les points $a \wedge b$ et $a \vee b$ y seront toujours (7) et (8).

Ad 1′. Soit c un point qui se trouve entre deux points a et b et soit x un point qui est plus prochain que a et b. Il est à montrer que x est aussi plus prochain que c. Autrement dit, soit

$$ab \subset c \subset a + b,$$
$$ua \subset x \subset u + a,$$
$$ub \subset x \subset u + b.$$

Il est à montrer que

$$uc \subset x \subset u + c.$$

Or, ceci résulte des relations:

$$uc \subset u(a + b) = ua + ub \subset x \subset (u + a)(u + b) = u + ab \subset u + c.$$

Ad 2′. Soit c un point qui se trouve entre deux points a et b et soit y un point qui est plus lointain que a et b. Il est à montrer que y est aussi plus lointain que c. Autrement dit, soit

$$ab \subset c \subset a + b,$$
$$uy \subset a \subset u + y,$$
$$uy \subset b \subset u + y.$$

Il est à montrer que

$$uy \subset c \subset u + y.$$

Or, ceci résulte des relations:

$$uy \subset ab \subset c \subset a + b \subset u + y.$$

Ad $a \wedge b$. Posons

$$x = ab + u(a + b) = ab + ua + ub.$$

Alors, on a, tout d'abord,

$$x = (a + b)(u + ab) = (a + b)(u + a)(u + b).$$

C'est une conséquence immédiate de (D_1). Puis, x se trouve entre a et b. Pour s'en convaincre, il suffit de remarquer que c'est (E_1) qui y est remplie, car on a, en tenant compte de (D_1),

$$ax + bx = a(ab + u(a + b)) + b(ab + u(a + b)) = ab + ua + ub = x$$

et

$$(a + x)(b + x) = (a + (a + b)(u + ab))(b + (a + b)(u + ab))$$
$$= (a + b)(u + a)(u + b) = x.$$

Soit enfin c un autre point qui se trouve entre a et b. Il est à montrer que x est plus prochain que c. Autrement dit, soit

$$ab \subset c \subset a + b.$$

Il est à montrer que

$$uc \subset x \subset u + c.$$

Or, ceci résulte des relations

$$uc \subset u(a + b) \subset ab + u(a + b) = x = (a + b)(u + ab) \subset u + ab \subset u + c.$$

Ad $a \vee b$. Posons

$$y = ab + \tilde{u}(a + b) = ab + \tilde{u}a + \tilde{u}b.$$

Alors on a, tout d'abord,

$$y = (a + b)(\tilde{u} + ab) = (a + b)(\tilde{u} + a)(\tilde{u} + b).$$

Puis, y se trouve entre a et b. Tout cela se démontre de même que les choses analogues pour x, en remplaçant x par y et u par $\tilde{u}$. Soit enfin c un autre point qui se trouve entre a et b. Il est à montrer que y est plus lointain que c. Ce dernier signifie que c se trouve entre u et y. Or, c'est équivalent, en vertu du théorème IV, à ce que y se trouve entre c et $\tilde{u}$. Autrement dit, soit

$$ab \subset c \subset a + b.$$

Il est à montrer que

$$\tilde{u}c \subset y \subset \tilde{u} + c.$$

Celui-ci se démontre de même que la chose analogue pour x, en remplaçant toujours x par y et u par $\tilde{u}$.

Nous avons démontré ainsi la suffisance de la condition du théorème.

Soit maintenant S un système métrisé, de choses complètement normées, et soit u un point, de l'espace formé par S, qui peut être pris pour l'origine de cet espace. Nous allons voir que u doit être une extrémité de S, c'est-à-dire qu'il doit posséder les propriétés 1) et 2) (n° 5).

En vertu de la propriété 2′ de l'origine, le point u ne peut être choisi pour l'origine que si, après ce choix, parmi les points qui se trouvent entre les deux points arbitraires a et b, il existera un qui sera le plus prochain et il existera un autre qui sera le plus lointain. Or, tous les points de l'espace formé par S

se trouvent entre les deux points déterminés, savoir entre 0 et 1, car on a, pour chaque chose z, de S,

$$0z + 1z = z = (0 + z)(1 + z),$$

c'est à dire que (E_1) y est remplie. Donc, on peut affirmer, en particulier, que, parmi *tous* les points de l'espace formé par S, il existera un qui sera le plus lointain. Soit v ce point. En particulier, le point v sera plus lointain que le point 0 lui-même, c'est-à-dire que 0 se trouvera entre u et v, ou bien que

$$u0 + v0 = 0 = (u + 0)(v + 0),$$

ou encore

$$uv = 0. \tag{9}$$

De même, le point v sera plus lointain que le point 1, c'est-à-dire que 1 se trouvera, lui-aussi, entre u et v, ou bien que

$$u1 + v1 = 1 = (u + 1)(v + 1),$$

ou encore

$$u + v = 1. \tag{10}$$

Revenons aux points quelconques a et b. Dire que, parmi les points c qui se trouvent entre a et b, il existera un qui est le plus prochain, c'est dire qu'il existera un point x tel que, premièrement, x se trouve entre a et b et, deuxièmement, quel que soit le point c se trouvant entre a et b, x se trouve entre u et c. Alors, puisque les points ab et $a + b$ se trouvent toujours entre a et b, il existera, en particulier, un point x qui se trouve, en même temps, entre a et b, entre u et ab et entre u et $a + b$, de sorte qu'on a

$$ab \subset x \subset a + b,$$
$$uab \subset x \subset u + ab,$$
$$u(a + b) \subset x \subset u + (a + b).$$

Nous allons voir, tout d'abord, que le seul point x pouvant satisfaire à ces trois conditions est

$$x = ab + u(a + b) = (a + b)(u + ab).$$

En effet, on doit y avoir

$$ab + u(a + b) \subset x \subset (a + b)(u + ab).$$

Par suite, dire qu'il existera un point qui se trouve entre a et b et qui satisfait à trois conditions en question, c'est dire, en particulier, que le point $ab + u(a + b)$ se trouvera entre a et b. Or, dans l'article cité, j'ai établi [4] que, lorsqu'on a

[4] *Au cours de la démonstration du théorème VI*, pp. 821-823.

$$(a+b)u \neq au + bu,$$

alors le point $ab + u(a+b)$ ne se trouve pas entre a et b. Par suite, dire que le point $ab + u(a+b)$ se trouvera entre a et b, c'est dire, en particulier, qu'on aura

$$(11) \qquad (a+b)u = au + bu.$$

Pareillement, dire que, parmi les points c qui se trouvent entre a et b, il existera un qui est le plus lointain, c'est-dire qu'il existera un point y tel que, premièrement, y se trouve entre a et b et, deuxièmement, quelque soit le point c se trouvant entre a et b, c se trouve aussi entre u et y. Or, ce dernier revient au dire que y se trouve entre c et v. En effet le point v étant plus lointain que tout autre point, les points c et y se trouvent, tous les deux, entre u et v, de sorte qu'on a

$$(u,c) + (c,v) = (u,v),$$
$$(u,y) + (y,v) = (u,v).$$

On en tire:

$$(u,c) + (c,v) = (u,y) + (y,v).$$

De plus, lorsque c se trouve entre u et y, on a

$$(u,c) + (c,y) = (u,y).$$

En comparant, on obtient:

$$(c,y) + (y,v) = (c,v).$$

Cela établi, on peut affirmer qu'on aura:

$$(12) \qquad (a+b)v = av + bv.$$

On la démontre de même que la chose analogue pour x, en remplaçant x par y et u par v.

En résumé, on voit qu'à la chose u doit correspondre une chose v, de sorte que u et v satisfassent aux égalités (9), (10), (11) et (12). Or, ces dernières expriment précisement les propriétés 1) et 2) caractérisant l'extrémité u et son opposée $\tilde{u}$.

Nous avons démontré ainsi la nécessité de la condition du théorème.

8. THÉORÈME VI. *Si, dans l'espace formé par un système métrisé S, de choses complètement normées, tout point de l'espace peut être pris pour l'origine, le système S est nécessairement distributif.*

C'est un simple corollaire du Théorème V. En effet, lorsque tout point c, de l'espace formé par S, peut être pris pour l'origine, alors, en vertu du

théorème V, toute chose c, de S, est une extrémité de S. Dans ce cas, quelles que soient les choses a, b, c, de S, on doit avoir

$$(a + b)c = ac + bc.$$

C'est la loi distributive.

V. Cas des systèmes distributifs.

9. La condition T, nécessaire et suffisante pour que l'espace métrique presque ordonné soit formé par un système distributif, reste intacte quand on varie le choix de l'origine de cet espace. En effet, la condition T ne contient aucune indication, soit explicite soit implicite, sur le choix de l'origine.

Il en est autrement pour une autre condition, T bis, établie dans mon article cité. Pour bien comprendre tout ce qui suit, rappelons d'abord un théorème auxiliare qui y était établi, à savoir:

Lorsque, dans l'espace métrique formé par un système de choses normées, on prend 0 pour l'origine, un point a est plus prochain qu'un point b si et seulement si l'on a

$$a \subset b.$$

Après avoir démontré ce théorème, j'ai commencé d'employer le symbole $a \subset b$, non pas seulement dans son sens primitif, celui d'une relation entre les deux choses a et b d'un système, mais aussi pour exprimer que le point a, de l'espace formé sur ce système, est plus prochain que le point b. Pareillement, il y était établi que, lorsqu'on prend 0 pour l'origine, on a $a \wedge b = ab$ et $a \vee b = a + b$. Après ceci, j'ai commencé d'employer les symboles ab et $a + b$ pour désigner les points $a \wedge b$ et $a \vee b$. C'est ce qu'on doit avoir en vue dans l'énoncé du théorème concernant la condition T bis, à savoir:

La condition nécessaire et suffisante pour qu'un espace métrique presque ordonné D soit transitif est qu'il possède la propriété suivante:

T bis. *Un point c, de D, se trouve entre deux points a et b si et seulement si l'on a*

$$ab \subset c \subset a + b.$$

On y suppose, comme je le faisais partout dans l'article cité, que c'est 0 qui y est pris pour l'origine. Donc, en tenant compte des remarques qui viennent d'être faites, l'expression $ab \subset c \subset a + b$ qui figure dans l'énoncé de ce dernier théorème peut s'interpréter de deux façons différentes. Nous allons exprimer cela explicitement en forme de deux théorèmes, VII et VIII. Ce qui est intéressant, c'est que ces théorèmes-ci restent valables avec le choix arbitraire de l'origine.

THÉORÈME VII. *La condition nécessaire et suffisante pour qu'un espace métrique presque ordonné D formé par un système S de choses normées, soit transitif est qu'il possède la propriété suivante:*

T bis. *Un point c, de D, se trouve entre deux points a et b si et seulement si l'on a, pour les choses a, b, c de S,*

$$ab \subset c \subset a + b.$$

La démonstration y est inutile, car c'est précisemment ce qui a été demontré dans mon article cité.

THÉORÈME VIII. *La condition nécessaire et suffisante pour qu'un espace métrique presque ordonné D soit transitif est qu'il possède la propriété suivante:*

T ter. *Un point c, de D, se trouve entre deux points a et b si et seulement si ce point c est plus lointain que $a \wedge b$ et plus prochain que $a \vee b$.*

Démonstration. Supposons d'abord que l'espace en question est transitif, c'est-à-dire que la condition T y est remplie, et soit c un point qui est plus lointain que $a \wedge b$ et plus prochain que $a \vee b$. Nous allons voir que c se trouve alors entre $a \wedge b$ et $a \vee b$. En effet, d'après les suppositions faites, le point $a \wedge b$ se trouve entre l'origine u et le point c, tandis que le point c se trouve entre l'origine u et le point $a \vee b$. Ce dernier peut s'exprimer aussi, en vertu du Théorème IV, en disant que le point $a \vee b$ se trouve entre le point c et le point $\tilde{u}$ qui est l'extrémité opposée de u. On a donc, en vertu du Théorème II,

$$uc \subset a \qquad b \subset u + c,$$

$$\tilde{u}c \subset a \vee b \subset \tilde{u} + c.$$

De plus, comme, en vertu du Théorème III, *tous* les points se trouvent entre u et $\tilde{u}$, on a, d'après (E_1),

$$uc + \tilde{u}c = c = (u + c)(\tilde{u} + c).$$

On en déduit sans peine qu'on a

$$(a \wedge b + c)(a \vee b) \subset (u + c)(\tilde{u} + c) = c = uc + \tilde{u}c \subset (a \wedge b)c + (a \vee b).$$

En tenant compte de (E_2), ceci suffit pour affirmer que le point c se trouve entre les points $a \wedge b$ et $a \vee b$. Or, ces derniers points se trouvent entre a et b. Par suite, d'après T, le point c se trouve, lui-aussi, entre a et b. On voit ainsi que, si le point c est plus lointain que $a \wedge b$ et plus prochain que $a \vee b$, ce point c se trouve entre a et b. La réciproque étant manifeste, cela signifie que la condition T ter est remplie.

Inversement, supposons que la condition T ter soit remplie, que c se trouve entre x et y et que x et y se trouvent entre a et b. Alors, nous savons qu'on a

$$xy \subset c \subset x + y,$$
$$ab \subset x \subset a + b,$$
$$ab \subset y \subset a + b,$$

d'où

$$ab \subset c \subset a + b.$$

Ceci posé, on peut démontrer que le point c est plus lointain que $a \wedge b$ et plus prochain que $a \vee b$, c'est à dire que

$$uc \subset a \wedge b \subset u + c,$$
$$\tilde{u}c \subset a \vee b \subset \tilde{u} + c,$$

ou bien que

$$uc \subset ab + u(a + b) \subset u + c,$$
$$\tilde{u}c \subset ab + \tilde{u}(a + b) \subset \tilde{u} + c.$$

Pour s'en convaincre, il suffit de remarquer que, lorsqu'on a $ab \subset c$, on a aussi

$$ab + u(a + b) \subset c + u(a + b) \subset c + u,$$
$$ab + \tilde{u}(a + b) \subset c + \tilde{u}(a + b) \subset c + \tilde{u},$$

et que, lorsqu'on a $c \subset a + b$, on a aussi

$$uc \subset u(a + b) \subset ab + u(a + b),$$
$$\tilde{u}c \subset \tilde{u}(a + b) \subset ab + \tilde{u}(a + b).$$

Ceci établi, il en resulte, d'après T ter, que c se trouve entre a et b. On voit ainsi que, si le point c se trouve entre x et y et que x et y se trouvent entre a et b, le point c se trouve, lui-aussi, entre a et b. C'est la condition T, de sorte que l'espace est transitif.

10. Dans le cas des systèmes distributifs, l'énoncé du théorème fondamental se simplifie comme il suit.

THÉORÈME IX. *Dans l'espace formé par un système métrisé distributif S, de choses complètement normées, un point u peut être pris pour l'origine si et seulement si la chose u, de S, possède le complément, c'est à dire qu'il existe la chose $\tilde{u}$ liée avec u par les relations*

$$u\tilde{u} = 0,$$
$$u + \tilde{u} = 1.$$

C'est une conséquence immédiate du Théorème V. En effet, la propriété de u qui figure dans l'énoncé du présent théorème n'est autre chose que la

propriété 1) de l'extrémité. Quant à la propriété 2) de l'extrémité, elle est remplie, dans les systèmes distributifs, de soi-même.

11. Il est intéressant de considérer encore le cas particulier des systèmes *ordonnés*, c'est-à-dire des systèmes où, pour chaque couple de choses a, b, on a nécessairement, ou bien $a \subset b$ ou bien $b \subset a$.

THÉORÈME X. *Dans l'espace formé par un système métrisé ordonné S, de choses complètement normées, les seuls points qui peuvent être pris pour l'origine sont les point* 0 *et* 1.

En effet, l'origine devant être, d'après le théorème V, une extrémité de S, soit u cette extrémité et soit $\tilde{u}$ son opposée. Alors, si l'on a $u \subset \tilde{u}$, on a $u = u\tilde{u} = 0$. Si l'on a, au contraire, $\tilde{u} \subset u$, on a $u = u + \tilde{u} = 1$.

VI. Un exemple.

12. Terminons par un exemple. Soit S l'ensemble dont les éléments $a, b, \cdots$ sont les fonctions mesurables $\phi(\xi)$, d'une variable réelle ξ, définies dans l'intervalle $0 \leq \xi \leq 1$ et telles que $0 \leq \phi(\xi) \leq 1$. Convenons d'écrire $a = b$ si les fonctions correspondantes ne diffèrent l'une de l'autre que sur un ensemble de mesure nulle, et posons $a \subset b$ si, pour les fonctions correspondantes $\phi(\xi)$ et $\psi(\xi)$, on a presque partout $\phi(\xi) \leq \psi(\xi)$. Enfin, a étant la fonction $\phi(\xi)$, posons

$$| a | = \int_0^1 \phi(\xi)\, d\xi.$$

Cela posé, l'ensemble S est un système distributif de choses complètement normées. Ici, on a, a et b étant respectivement les fonctions $\phi(\xi)$ et $\psi(\xi)$,

$$ab = \min[\phi(\xi), \psi(\xi)], \qquad a + b = \max[\phi(\xi), \psi(\xi)].$$

Quant aux extrémités opposées, u et $\tilde{u}$, ce sont respectivement les fonctions $\omega(\xi)$ et $\tilde{\omega}(\xi)$ de la forme suivante. Soient Ξ_1 et Ξ_2 deux ensembles mesurables, de points ξ, dont la partie commune est vide et dont la somme est l'intervalle $0 \leq \xi \leq 1$, d'ailleurs arbitraires. Alors

$$\omega(\xi) = \begin{cases} 0 \text{ sur } \Xi_1, \\ 1 \text{ sur } \Xi_2, \end{cases} \qquad \tilde{\omega}(\xi) = \begin{cases} 1 \text{ sur } \Xi_1, \\ 0 \text{ sur } \Xi_2. \end{cases}$$

Dans l'espace formé par le système S métrisé, la distance de point a et de point b, qui sont respectivement les fonctions $\phi(\xi)$ et $\psi(\xi)$, est donnée par la formule

$$(a, b) = \int_0^1 | \phi(\xi) - \psi(\xi) |\, d\xi.$$

Ici, un point c, qui est la fonction $\chi(\xi)$, se trouve entre deux points a et b, qui sont respectivement les fonctions $\phi(\xi)$ et $\psi(\xi)$, si et seulement si l'on a presque partout

$$\min[\phi(\xi), \psi(\xi)] \leq \chi(\xi) \leq \max[\phi(\xi), \psi(\xi)].$$

Pour l'origine, on peut y prendre un point u qui est l'une des fonctions $\omega(\xi)$ définies plus haut. Enfin, a et b étant respectivement les fonctions $\phi(\xi)$ et $\psi(\xi)$, on a

$$a \wedge b = \begin{cases} \min\ [\phi(\xi), \psi(\xi)] & \text{sur } \Xi_1, \\ \max\ [\phi(\xi), \psi(\xi)] & \text{sur } \Xi_2, \end{cases}$$

$$a \vee b = \begin{cases} \max\ [\phi(\xi), \psi(\xi)] & \text{sur } \Xi_1, \\ \min\ [\phi(\xi), \psi(\xi)] & \text{sur } \Xi_2. \end{cases}$$

Appendice.

Je profite de l'occasion pour corriger une petite erreur admise dans mon article cité. On lit, p. 804: "Pour avoir un exemple d'espace métrique presque ordonné, prenons un ensemble arbitraire de nombres réels, où $(a, b) = |a - b|$. Il est aisé de voir que le rôle de l'origine peut être joué par un nombre quelconque appartenant à cet ensemble." Le Théorème X du présent article nous apprend que cet exemple est incorrect.

Cette même erreur se répète p. 818, où l'on lit: "Pour avoir un exemple d'espace métrique presque ordonné transitif, il suffit de rappeler l'exemple déjà cité d'un ensemble arbitraire de nombres réels, où $(x, y) = |x - y|$."

Moscou, U. S. S. R.

TOPOLOGICAL PROPERTIES OF THE SOLUTIONS OF ORDINARY DIFFERENTIAL EQUATIONS.*

By T. M. Cherry.

The conception that any solution of a set of differential equations possesses topological properties rests upon its representation as a curve in a suitably chosen space, and the observation that the points of this curve do not in general form a closed set. Hence we are led to consider its derived set, which consists of the curve itself together with two other sets, its ω- and α-limit sets, which are derived respectively from the "positive and negative ends" of the curve (Birkhoff [1]). The present paper rests upon the idea that solutions may be classified according to their relations with the associated ω- and α-limit sets. This idea is not new, the case of a solution which is contained by both its limit sets having been long recognized as of importance (Hadamard [1], Birkhoff and Smith [1]); but it does not appear hitherto to have been systematically developed. Such a development is found in §§ 4, 5, 7 below, in relation to which §§ 1-3 contain the necessary preliminaries and § 6 is a digression. The recurrent solutions of Birkhoff appear in § 5 as a sub-class of those I call pervasive, while the theory of general pervasive solutions in § 7 amounts to a generalization of the descriptive theory of transitive systems. In §§ 8-11 there is developed a method of constructing transitive systems which do not consist simply of a single minimal set, and yet are soluble by quadratures; the systems constructed are not however of Hamiltonian type.

The extent of previous explorations in the field with which we are concerned makes it inevitable that the present work should be distinctive in its point of view rather than its method or final results. Theorems II, IV, V and part of Theorem VI are apparently new, while the proofs of Theorems III (i), VI, VII differ more or less in detail from those given by Birkhoff. The examples in §§ 9-11 are new.

[*Note added in proof.* A Referee has drawn my attention to a recent paper of H. Hilmy [1], the ideas of which are to some extent common with those here used, though differently introduced. Hilmy's *exceptional undecomposable set* is in my terminology (§ 4) an asymptotic trajectory together with its ω- and α-limit sets; his *quasi-minimal set* is my *stratum* (§ 4); and

---

* Received March 4, 1937.

his example of a quasi-minimal set which is not minimal anticipates the principle used in § 9.]

1. Terminology. Let

$$dx_i/dt = X_i(x_1, \cdots, x_n), \qquad (i = 1, \cdots, n) \tag{1}$$

be a system of differential equations in which $X_1, \cdots, X_n$ are real single-valued functions of n real dependent variables $x_1, \cdots, x_n$. The conception that is fundamental to our object is that any set of real values $(x_1, \cdots, x_n)$ of these variables is a single entity, a *point* of an n-dimensional space; so that by suitably defining one or both of the terms *neighbourhood* and *distance* for this space we can utilize the topological ideas of limit-point, etc. It is customary to take the space to be Euclidean, but that this convention is unnecessarily restrictive is evident since the results to be proved are purely topological. We shall take as our fundamental definition: *A neighbourhood of a point* (a_i) *is the aggregate of all points* (x_i) *for which, for some positive number* ϵ,

$$| x_i - a_i | < \epsilon \qquad (i = 1, \cdots, n).$$

More specifically, we call this aggregate *the* ϵ*-neighbourhood of* (a_i). We utilize also *closed neighbourhoods* defined by inequalities $| x_i - a_i | \leq \epsilon$.

For convenience of language we abridge certain statements about neighbourhoods by use of the term *distance*, with the definition: *The distance between the points* (x_i), (a_i) *is the greatest of the n numbers* $| x_i - a_i |$. The symmetry and triangle axioms are thereby satisfied; and the statement "the distance between (x_i), (a_i) is equal to (less than, greater than) ϵ" is equivalent to "each of the points (x_i), (a_i) lies on the boundary of (inside, outside) the ϵ-neighbourhood of the other."

We describe the functions X_i as *regular* at a point (a_i) if *they all have continuous first derivatives in some neighbourhood of the point.* If the X_i are regular at (a_i) a solution

$$x_i = f_i(t) \qquad (i = 1, \cdots, n) \tag{2}$$

of the system (1), for which $f_i(t_0) = a_i$, is uniquely defined in some interval $| t - t_0 | < \tau$; and the solution may be continued over successive intervals overlapping with this as long as the point $(f_i(t))$ does not approach arbitrarily near to a point where the X_i are not all regular.

We describe the solution (2) as *regular* if *the* $f_i(t)$ *may be continued over all real values of* t, *and if there is a positive constant* δ *such that the* X_i *are regular in the* δ*-neighbourhood of each of its points.* We describe the solution

as *bounded* if *the* $f_i(t)$ *are bounded.* The existence of regular bounded solutions [1] is assured if the system (1) possesses an integral

$$F(x_1, \cdots, x_n) = \text{constant} = c$$

which, for some range of c, represents in our space a family of hyper-surfaces possessing closed sheets, at all points of which the X_i are regular. Our object being to investigate properties of a solution which follow merely from the hypothesis that it is bounded and regular, it would be beside the point here to develop other criteria for the existence of such solutions.

Any solution (2) specifies in our space a variable point $(f_i(t))$ which we call a *current point* and denote by a symbol such as $P(t)$; definite positions of it (fixed points of space) are denoted by symbols such as $P(t_0), P_0, P$; the aggregate of all its positions for $-\infty < t < \infty$ is a solution-curve or *trajectory.* Any *solution* specifies a trajectory together with a correspondence between its points and values of t. It is often (e. g. in the phrase "a point of a solution [trajectory]") but not always a matter of indifference whether we speak of "solution" or "trajectory"; and we use the same symbol, such as (P), for both. We call t the time-variable, and use terms such as *time, instant, prior* in connection therewith; and we can hence think of a solution as being or specifying a *motion* of a current point.

From the solution (2) we can derive another

$$x_i = f_i(t + \tau) \qquad (i = 1, \cdots, n) \tag{3}$$

in which τ is any constant. To the solutions (2), (3) correspond different current points $P(t)$, $P(t + \tau)$ which run along the same trajectory, the former passing through any position a time τ in advance of the latter. Thus given any two positions P_0, P_1 on the trajectory, all current points thereon take the same time, T say, to pass from P_0 to P_1, and the time $(-T)$ to pass from P_1 to P_0. We call P_1 the *T-transform* [2] of P_0, and P_0 the $(-T)$-transform of P_1.

2. The auxiliary theorem. Let $\mathfrak{D}$ be a closed connected set of points, and $\mathfrak{D}'$ the (open) domain which is the sum of the δ-neighbourhoods of all points of $\mathfrak{D}$; δ being a positive constant. Let the X_i be regular at all points of $\mathfrak{D}'$. Then, from the fundamental theorems on the existence of solutions and their continuity with respect to the initial values, there follows:

---

[1] Birkhoff [1] calls such solutions *positively and negatively stable.*

[2] Poincaré [1] calls P_1 a *consequent* or *antecedent* of P_0 according as T is positive or negative.

(i) *There is a positive constant* τ *such that any current point which is in* $\mathfrak{D}$ *for* $t = t_0$ *remains within* $\mathfrak{D}'$ *throughout the interval* $|t - t_0| < \tau$ *at least.*

(ii) *Let a current point* $P(t)$ *be at* $P(t_0)$ *in* $\mathfrak{D}$ *for* $t = t_0$, *and remain within* $\mathfrak{D}'$ *for* $t_1 < t < t_2$, (*where* t_1 *may be* $-\infty$ *and* t_2 *may be* ∞); *let* ϵ *be any positive number and* T *any* (*finite*) *number between* t_1 *and* t_2; *then there exists a positive number* η, $\equiv \eta(\epsilon, T)$, *such that all current points whose distance from* $P(t_0)$ *at* $t = t_0$ *does not exceed* η *are at a distance from* $P(t)$ *not exceeding* ϵ *for all* t *in the closed interval* (t_0, T).

Following Birkhoff [1], we call (ii) the Auxiliary Theorem.

COROLLARY. *If* $P(t_0)$ *is a limit-point* (*or the sole limit-point*) *of a sequence of points, then* $P(t_0 + T)$ *is a limit-point* (*or the sole limit-point*) *of the* T*-transform of this sequence, provided* $t_1 < t_0 + T < t_2$.

3. The ω- and α-limit sets of a bounded regular solution. The following definitions and theorems are due, in their precise form, to Birkhoff [1].

Let (P) be a bounded regular solution and $P(t)$ the associated current point. Let $\{t_i\}$ be any increasing unbounded sequence. The hypothesis that the solution is bounded implies that the points $P(t_i)$ all lie in a bounded portion of the space; hence the set $\{P(t_i)\}$ has at least one limit point, called an *ω-limit point of* (P). The set of all such points is called the *ω-limit set of* (P), and denoted by $L_\omega(P)$.

The set of points similarly defined by decreasing unbounded sequences $\{t_i\}$ is called the *α-limit set of* (P), and denoted by $L_\alpha(P)$.

It is apparent that all solutions which, like (2) and (3), belong to the same trajectory define the same ω- and α-limit sets.

The derived set of (P), regarded simply as a set of points, consists of $L_\omega(P)$, $L_\alpha(P)$ together with the set (P) itself, which may or may not belong to either or both of $L_\omega(P)$, $L_\alpha(P)$.

THEOREM I. *The following results are stated for the ω-limit set; there are corresponding results for the α-limit set.*

(i) *The set* $L_\omega(P)$ *is closed,* since in an arbitrary neighbourhood of any point of its derived set there are points of $L_\omega(P)$ and hence points $P(t)$ for which t is arbitrarily large.

(ii) *The set* $L_\omega(P)$ *consists of a set of bounded regular trajectories.* If Q is any point of $L_\omega(P)$ we are to prove that for any T the T-transform of Q belongs to $L_\omega(P)$.

Call $\mathcal{D}'$ the δ-neighbourhood of the trajectory (P), i. e. the sum of the δ-neighbourhoods of all its points: δ being the constant involved by definition (§ 1) in the hypothesis that (P) is regular. By the Auxiliary Theorem, if Q_1, the T-transform of Q, belongs to $\mathcal{D}'$, and if Q is a limit-point of the sequence $\{P(t_i)\}$ in which $t_i \to \infty$ as $i \to \infty$, then Q_1 is a limit-point of the sequence $\{P(t_i + T)\}$, and is an ω-limit-point since $t_i + T \to \infty$ as $i \to \infty$. Hence *if the T-transform of Q belongs to $\mathcal{D}'$ it belongs to $L_\omega(P)$.*

But by § 2 (i) there is a positive constant τ such that, if $Q(t')$ belongs to $L_\omega(P)$, (every point of which is within $\mathcal{D}'$ at a distance at least δ from its boundary), then $Q(t)$ belongs to $\mathcal{D}'$ for $|t - t'| < \tau$; so by the result just italicized, if $Q(t)$ belongs to $L_\omega(P)$ for $t' \leq t \leq t''$, it belongs to $L_\omega(P)$ for $t' - \tau < t < t + \tau$. Hence $Q(t)$ belongs to $L_\omega(P)$ for all real t; which proves the result.

It is clearly appropriate to say that $L_\omega(P)$ consists of a closed set of trajectories or solutions.

(iii) *Corresponding to any positive number ϵ there exists a (finite) number $T, \equiv T(\epsilon)$, such that for $t > T$, $P(t)$ lies within the ϵ-neighbourhood of $L_\omega(P)$.* This is proved by reductio ad absurdum. In virtue of the result we can say that, if (P) is not a part of $L_\omega(P)$, $P(t)$ is *asymptotic* to $L_\omega(P)$ as $t \to \infty$.

(iv) From (iii) it follows that *the set $L_\omega(P)$ is connected*; and hence from (i) *it is perfect.*

(v) *To any numbers t_0 and ϵ, $(\epsilon > 0)$, corresponds a (finite) number $T', \equiv T'(\epsilon, t_0)$, greater than t_0 such that, for some t between t_0 and T', $P(t)$ is within the ϵ-neighbourhood of each point of $L_\omega(P)$*; i. e. $P(t)$ approaches its limit-points $L_\omega(P)$ uniformly. This is proved by reductio ad absurdum.

4. Classification of bounded regular solutions. It follows from § 3 (ii) that, if one point of the solution (P) belongs to $L_\omega(P)$, then all its points belong. Hence there are two mutually exclusive possibilities, which we denote by the symbols V_ω, A_ω:

V_ω : every point of (P) belongs to $L_\omega(P)$,
A_ω : no point of (P) belongs to $L_\omega(P)$.

Similarly there are two mutually exclusive possibilities, which are shewn by particular examples (§§ 9, 10 below) to be independent of the former:

V_α : every point of (P) belongs to $L_\alpha(P)$,
A_α : no point of (P) belongs to $L_\alpha(P)$.

By combination of these we classify bounded regular solutions (P) into three types, as follows:

PERVASIVE SOLUTIONS $(V_\omega V_\alpha)$. (P) *belongs to both* $L_\omega(P)$ *and* $L_\alpha(P)$.

ASYMPTOTIC SOLUTIONS $(A_\omega A_\alpha)$. (P) *belongs to neither* $L_\omega(P)$ *nor* $L_\alpha(P)$.

PERVASIVE-ASYMPTOTIC SOLUTIONS ($V_\omega A_\alpha$ OR $V_\alpha A_\omega$). (P) *belongs to one but not both of* $L_\omega(P)$, $L_\alpha(P)$.

Any solution which possesses the property V_ω will be called *positively semi-pervasive*; and similarly for the properties V_α, A_ω, A_α. Thus a solution of type $V_\alpha A_\omega$ is negatively semi-pervasive and positively semi-asymptotic.

It is apparent that all solutions which differ merely by constants additive with t are of the same type; and the trajectory which carries a pervasive (asymptotic, etc.) solution may be described as pervasive (asymptotic, etc.).

If a solution (P) is positively semi-pervasive the set of points $L_\omega(P)$ will be called the associated *stratum*;[3] and similarly for a negatively semi-pervasive solution. If the solution is pervasive it will be shewn below (Theorem II) that the sets $L_\omega(P)$, $L_\alpha(P)$ coincide, so that there is only one stratum associated with the solution.

A "pervasive" solution is so named because its points are everywhere dense in, or pervade, the associated stratum. The term "recurrent" would be equally appropriate, but is already standardized as denoting a sub-class of pervasive solutions, namely, uniformly pervasive solutions (§ 5 below). The terms stable-à-la-Poisson (Poincaré [2]) quasi-periodic (Husson [1]) and quasi-recurrent (Birkhoff and Smith [1]) have also been used with the same sense. An "asymptotic" solution is named from the property pointed out in § 3 (iii) (*cf.* Hadamard [1]).

In the following work the symbol $\subset$ is used to denote the relation of *strict* inclusion between the two sets; so that $S_1 \subset S_2$ excludes the possibility $S_1 = S_2$.

THEOREM II. *For a pervasive solution* (P), $L_\omega(P) = L_\alpha(P)$. *For a pervasive-asymptotic solution of type* $V_\omega A_\alpha$, $L_\alpha(P) \subset L_\omega(P)$. *For a pervasive-asymptotic solution of type* $V_\alpha A_\omega$, $L_\omega(P) \subset L_\alpha(P)$.

Let us assume the hypothesis V_ω, i. e. that (P) is positively semi-pervasive. Every α-limit point of (P) is a limit point of a sequence $\{P(t_i)\}$ in which

[3] Hadamard [1] calls this the *domain of* (P); as this usage of the word "domain" is not in accord with its accepted modern connotation I introduce a new term.

$\{t_i\}$ decreases to $-\infty$. Since all the points of this sequence belong to (P), and hence by hypothesis to $L_\omega(P)$, and the latter set is closed, every α-limit point of (P) belongs to $L_\omega(P)$. Hence

(4) *if* $(P) \subseteqq L_\omega(P)$, *then* $L_\alpha(P) \subseteqq L_\omega(P)$.

Similarly, *if* $(P) \subseteqq L_\alpha(P)$, *then* $L_\omega(P) \subseteqq L_\alpha(P)$.

Hence if (P) is pervasive, so that $(P) \subseteqq L_\omega(P)$ and $(P) \subseteqq L_\alpha(P)$, then $L_\alpha(P) = L_\omega(P)$.

Conversely, assume that $(P) \subseteqq L_\omega(P)$ and $L_\alpha(P) = L_\omega(P)$; then $(P) \subseteqq L_\alpha(P)$, and the solution (P) is pervasive.

Hence if $(P) \subseteqq L_\omega(P)$ and (P) is not pervasive, so that though positively semi-pervasive it is negatively semi-asymptotic, the possibility $L_\omega(P) = L_\alpha(P)$ is excluded, and from (4) there follows $L_\alpha(P) \subset L_\omega(P)$. Q. E. D.

It may be noted that for a solution of type $V_\omega A_\alpha$ we have $(P) \subset L_\omega(P)$ and not $(P) = L_\omega(P)$; for the latter alternative in combination with the Theorem just proved would give $L_\alpha(P) \subset (P)$, implying that some but not all points of (P) are α-limit points. If $(P) \subset L_\omega(P)$ the latter set consists of a non-enumerable set of trajectories, and if $(P) = L_\omega(P)$, (P) is periodic (Birkhoff [1]).

For an asymptotic solution there are no necessary relations of inclusion or exclusion between $L_\omega(P)$, $L_\alpha(P)$. Examples may be given of each of the cases (i) $L_\alpha(P) = L_\omega(P)$, (ii) $L_\omega(P)$, $L_\alpha(P)$ have no common point, (iii) one of $L_\omega(P)$, $L_\alpha(P)$ is a proper part of the other, (iv) $L_\omega(P)$, $L_\alpha(P)$ have common points, but neither contains the other.

The classical work of Poincaré [2] has often been quoted as proving, for the case in which there is a positive integral-invariant and all solutions which pass through a certain neighbourhood are bounded and regular, the existence of solutions of the types here called semi-pervasive and pervasive. This is not so; for Poincaré, taking an arbitrarily small but determinate (fixed) region $\mathfrak{D}$, proves the existence of solutions which traverse it infinitely often, and all that this shews is that the ω-limit set (or α-limit set, or both) traverse $\mathfrak{D}$. Poincaré's method is easily modified so as to yield the existence of pervasive solutions; but this is unnecessary since Carethéodory [1], who quotes Poincaré in the above incorrect sense, has proved that *the set of trajectories which are not positively (negatively) semi-pervasive is of measure zero*, and the complement of the sum of these two nul sets is the set of pervasive trajectories. Thus, with the hypothesis stated, *almost all trajectories are pervasive.*

5. Uniformly and non-uniformly pervasive solutions. The notion of a *uniformly pervasive* solution is due to Birkhoff [1], who calls it *recurrent*. From the present point of view it seems most natural to adopt as a definition the necessary and sufficient property of Birkhoff's Theorem III, instead of the notion of a minimal set, as follows:

Let (P) be a pervasive solution, and denote the associated stratum, i. e. the coincident sets $L_\omega(P)$, $L_\alpha(P)$, by $\mathfrak{S}$. Let us say that $P(t)$, starting from the position $P(t_0)$, makes an ϵ-circuit of $\mathfrak{S}$ in time T, $(T > 0)$, if for some t between t_0 and $t_0 + T$ (inclusive) it is in the closed ϵ-neighbourhood of each point of $\mathfrak{S}$. Then we define:

The pervasive solution (P) is uniformly pervasive or recurrent if for every positive number ϵ there is a (finite) time T, $\equiv T(\epsilon)$, within which a current point $P(t)$ makes an ϵ-circuit of $\mathfrak{S}$, whatever its starting point on (P).

THEOREM III.[4] (i) *If the solution (P) is uniformly pervasive its stratum $\mathfrak{S}$ consists of an aggregate of solutions each of which is uniformly pervasive with $\mathfrak{S}$ as its stratum.*

(ii) *If the solution (P) is non-uniformly pervasive, every solution which has one point in the stratum $\mathfrak{S}$ associated with (P) lies in $\mathfrak{S}$, as do its ω- and α-limit sets; and $\mathfrak{S}$ contains solutions for which both the ω- and α-limit sets are proper parts thereof.*

(i) Let any positive number ϵ be assigned. Choose any positive number ϵ' less that ϵ, let $\kappa = \epsilon - \epsilon'$, and let $T(\epsilon')$ be the time not exceeded, by definition, in any ϵ'-circuit of $\mathfrak{S}$ by $P(t)$.

Let Q be any point of $\mathfrak{S}$, $Q(t)$ the current point that is at Q at $t = 0$, and $R(t)$ any other current point. By the Auxiliary Theorem there is a positive η such that the distance $\overline{Q(t)R(t)}$ does not exceed κ for $0 \leq t \leq T(\epsilon')$ provided only the initial distance $\overline{QR(0)}$ does not exceed η. Take for $R(0)$ a point $P(t')$ whose distance from Q does not exceed η, a suitable choice of t' being possible since Q belongs to $\mathfrak{S}$. Then for $0 \leq t \leq T(\epsilon')$ the distance $\overline{Q(t)P(t'+t)}$ does not exceed κ. But in this time-interval $P(t'+t)$ passes within a distance not exceeding ϵ' of each point of $\mathfrak{S}$; so $Q(t)$ must pass within a distance not exceeding $\kappa + \epsilon', = \epsilon$; and as ϵ is arbitrary, every point of $\mathfrak{S}$ belongs to $L_\omega(Q)$.

But since $(Q) \subseteqq \mathfrak{S}$ and the latter set is closed, $L_\omega(Q) \subseteqq \mathfrak{S}$. Hence $L_\omega(Q) = \mathfrak{S}$. By a similar argument we prove $L_\alpha(Q) = \mathfrak{S}$; and since the

[4] This is identical with Birkhoff [1], Theorem III. The proof of (i) differs from Birkhoff's.

time for an ϵ-circuit by $Q(t)$ has been shewn not to exceed the constant $T(\epsilon')$, whatever its starting point, (Q) is uniformly pervasive.

(ii) The hypothesis that (P) is non-uniformly pervasive implies that, for some fixed positive ϵ, there are starting-points $P(t_n)$ from which $P(t)$ fails to make an ϵ-circuit of $\mathfrak{S}$ in the arbitrarily large time T_n. A limit point Q of $\{P(t_n + \frac{1}{2}T_n)\}$ for $T_n \to \infty$ gives, by use of the Auxiliary Theorem and § 3 (v), a solution (Q) for which neither $L_\omega(Q)$ nor $L_\alpha(Q)$ coincides with $\mathfrak{S}$. Since, as in (i), $L_\omega(Q)$ and $L_\alpha(Q)$ belong to $\mathfrak{S}$, we have $L_\omega(Q) \subset \mathfrak{S}$, $L_\alpha(Q) \subset \mathfrak{S}$.

THEOREM IV. *A uniformly semi-pervasive solution is uniformly pervasive.*

Suppose that for a positively semi-pervasive solution (P), $P(t)$ passes within an arbitrary distance ϵ of each point of $L_\omega(P)$ for $t_0 \leq t \leq t_0 + T(\epsilon)$, where $T(\epsilon)$ is finite and independent of t_0. Let Q be any point of $L_\omega(P)$. Choose any positive ϵ and a sequence $\{t_i\}$ diminishing to $-\infty$. Then for each i there is a t'_i between t_i and $t_i + T(\epsilon)$ for which $P(t'_i)$ is within a distance ϵ of Q. The sequence $\{P(t'_i)\}$ has a limit point R in the closed ϵ-neighbourhood of Q; and R belongs to $L_\alpha(P)$ since $t'_i \to -\infty$ as $i \to \infty$. Since ϵ is arbitrary and $L_\alpha(P)$ is closed, Q belongs to $L_\alpha(P)$. Hence $L_\omega(P) \subseteqq L_\alpha(P)$.

But since (P) is positively semi-pervasive, $L_\alpha(P) \subseteqq L_\omega(P)$, (Theorem II). Hence $L_\alpha(P) = L_\omega(P)$ and (P) is pervasive; and our initial assumption shews it to be uniformly pervasive.

6. A property of uniformly pervasive (recurrent) trajectories.

THEOREM V. *If $P(t)$, $P(t+T)$ are any two current points on the same uniformly pervasive trajectory, the distance $\overline{P(t)P(t+T)}$ has, for T fixed and t variable, a positive lower bound.*

If the trajectory (P) is periodic the result is trivial. In the contrary case let P_0, P_1 be the fixed non-coincident points $P(0)$, $P(T)$; then a domain $\mathfrak{D}_0$ surrounding P_0 has as its T-transform a domain $\mathfrak{D}_1$ surrounding P_1. If $\mathfrak{D}_0$ is infinitesimal so is $\mathfrak{D}_1$, so we can and shall suppose that the distance between points of $\mathfrak{D}_0$, $\mathfrak{D}_1$ has a positive lower bound δ.

Since (P) is pervasive, $P(t)$ is within $\mathfrak{D}_0$ at a sequence of instants

$$t = \cdots t_{-2}, t_{-1}, 0, t_1, t_2, \cdots, \tag{5}$$

where $t_i \to \infty$ and $t_{-i} \to -\infty$ as $i \to \infty$; and the points $P(t_i + T)$, being the T-transforms of $P(t_i)$, must be within $\mathfrak{D}_1$.

Now given any two distinct current points $Q(t)$, $R(t)$, not necessarily on the same trajectory, there is a *positive* lower bound for the distance $\overline{Q(t)R(t)}$ during a given finite time interval $t' \leq t \leq t' + T'$, and for continuous variation of the starting points $Q(t')$, $R(t')$ this bound varies continuously. Hence if $Q(t')$, $R(t')$ be any points of $\mathfrak{D}_0$, $\mathfrak{D}_1$ respectively, there is a positive lower bound g for $\overline{Q(t)R(t)}$ for $t' \leq t \leq t' + T'$.

Let ϵ be the upper bound of the distances of points of $\mathfrak{D}_0$ from P_0 and, introducing the hypothesis of *uniform* pervasiveness, let T_ϵ be the time in which $P(t)$ uniformly performs an ϵ-circuit of the associated stratum; and let τ be the upper bound to the time any current point remains continuously within $\mathfrak{D}_0$ (so that when ϵ is small, τ is small and T_ϵ large). Then in the sequence (5) we can, for every i, take $\tau < t_{i+1} - t_i \leq T_\epsilon + \tau$. Choose $T' = T_\epsilon + \tau$, $Q(t') = P(t_i)$ and $R(t') = P(t_i + T)$, and we have

$$\overline{P(t)P(t+T)} \geq g \quad \text{for} \quad t_i \leq t \leq t_{i+1}.$$

Since this holds for all integers i, the lower bound g is valid for all t.

Q. E. D.

For recurrent trajectories of elementary continuous type, specified by periodic functions of arguments $\lambda_1 t + a_1$, $\lambda_2 t + a_2, \cdots$ ($\lambda_1, \lambda_2 \cdots$ constants mutually incommensurable; $a_1, a_2 \cdots$ arbitrary constants), there are the further properties:

(i) The ratio of the upper to the lower bound of $\overline{P(t)P(t+T)}$ has, for all T, a finite upper bound; so that if by taking T suitably large we have $P(T)$ close to $P(0)$, the points $P(t+T)$, $P(t)$ remain permanently close.

(ii) The same is true of the bounds of $\overline{P(t)Q(t)}$, where $Q(t)$ is in the stratum associated with (P) (which is a minimal set) but not on the trajectory (P).

For recurrent trajectories in general it seems that these statements are false, and that it is not even true that $\overline{P(t)Q(t)}$ has a lower bound distinct from zero.

7. Solutions associated with a non-uniformly pervasive or pervasive-asymptotic solution. With any pervasive or semi-pervasive solution (P) is associated its stratum $\mathfrak{S}$, ($\mathfrak{S} = L_\omega(P)$ or $L_\alpha(P)$), which is a perfect set of points in which the points of (P) are everywhere dense, and which contains all points of any solution having one point therein. If Q is any point of $\mathfrak{S}$ we have therefore

$$L_\omega(Q) \subseteq \mathfrak{S}, \qquad L_\alpha(Q) \subseteq \mathfrak{S}.$$

We call Q

a *π-point* if $L_\omega(Q) = L_\alpha(Q) = \mathfrak{S}$,
a *σ-point* if $L_\omega(Q) \subset \mathfrak{S}$ and $L_\alpha(Q) \subset \mathfrak{S}$,
a *μ-point* if either $L_\omega(Q) = \mathfrak{S}$ and $L_\alpha(Q) \subset \mathfrak{S}$, or $L_\omega(Q) \subset \mathfrak{S}$ and $L_\alpha(Q) = \mathfrak{S}$.

Every point of $\mathfrak{S}$ belongs to one of these three classes; and all points of the same trajectory are of the same class. If (Q) consists of σ-points, $L_\omega(Q)$ and $L_\alpha(Q)$ consist of σ-points. If (Q) consists of μ-points, that one of $L_\omega(Q)$, $L_\alpha(Q)$ which is a part only of $\mathfrak{S}$ consists of σ-points. If $\mathfrak{S}$ contains one μ-point, the μ-points are everywhere dense therein.

If (P) is uniformly pervasive, $\mathfrak{S}$ consists entirely (Theorem III (i)) of π-points.

If (P) is non-uniformly pervasive, $\mathfrak{S}$ contains σ-points (Theorem III (ii)) as well as the π-points of (P). If (P) is pervasive-asymptotic, $\mathfrak{S}$ contains μ-points (those of (P)), and, as just pointed out, σ-points. In these cases the uniformly pervasive solution or solutions which $\mathfrak{S}$ must contain (Birkhoff [1], Theorem II) must consist of σ-points.

THEOREM VI. *The stratum associated with a non-periodic pervasive or pervasive-asymptotic solution contains an aggregate, having the cardinal number of the continuum, of pervasive trajectories which are everywhere dense therein, i. e. which consist of π-points.*

(i) Let (P) be a solution which is positively semi-pervasive, it being irrelevant whether negatively it is semi-pervasive or semi-asymptotic provided only it is not periodic. Then $L_\omega(P)$ is the associated stratum. Choose any point $P(t_0)$ of (P), this not being an equilibrium point since then (P) would be (degenerately) periodic. We can thus choose a neighbourhood $\mathfrak{N}$ of $P(t_0)$ and a positive constant τ such that any current point inside $\mathfrak{N}$ at $t = 0$ is outside $\mathfrak{N}$ at $t = \pm\tau$. Thus for any trajectory which traverses $\mathfrak{N}$, the portion within $\mathfrak{N}$ consists of a set of distinct arcs which is finite or enumerably infinite. In particular the trajectory (P) has infinitely many arcs within $\mathfrak{N}$; for since $P(t_0)$ belongs to $L_\omega(P)$, $P(t)$ must traverse $\mathfrak{N}$ infinitely often, and no two traverses are made on the same arc since (P) is not periodic.

(ii) Let us say that two closed neighbourhoods within $\mathfrak{N}$ are *not coupled* if there is no trajectory arc within $\mathfrak{N}$ having points common with both.

Let $\mathfrak{N}'$ be a neighbourhood of a point of (P), contained by $\mathfrak{N}$. Then $\mathfrak{N}'$ contains two closed neighbourhoods, centred at points of (P), which are not coupled. For since $\mathfrak{N}'$ is centred at a point of (P), which belongs to $L_\omega(P)$, $P(t)$ traverses $\mathfrak{N}'$ infinitely often on different arcs within $\mathfrak{N}$; the

closed ϵ-neighbourhoods of points P_0, P_1 of two such arcs, within $\mathfrak{N}'$, are then not coupled if ϵ is sufficiently small.

(iii) *Let there be assigned any closed neighbourhood $\bar{\mathfrak{N}}$ centred at any point of (P), and any positive numbers η, T; then there exists a neighbourhood $\phi(\bar{\mathfrak{N}}, \eta, T)$ centred at a point of (P), contained by $\bar{\mathfrak{N}}$, and such that any current point which is within $\phi(\bar{\mathfrak{N}}, \eta, T)$ at $t = 0$ is within the η-neighbourhood of $P(t_0)$ at two instants at least, one prior to $t = -T$ and the other subsequent to $t = T$.*[5] Let $P(t_1)$ be the centre and ϵ the radius of $\bar{\mathfrak{N}}$, i. e. $\bar{\mathfrak{N}}$ is the closed ϵ-neighbourhood of $P(t_1)$; and let $\epsilon' = \frac{1}{2}\epsilon$, $\eta' = \frac{1}{2}\eta$. Since $P(t_0)$ and $P(t_1)$ belong to $L_\omega(P)$ there exists a finite T' greater than T such that $P(t_0 + T')$ is within the ϵ'-neighbourhood of $P(t_1)$, and a finite T'' greater than T such that $P(t_0 + T' + T'')$ is within the η'-neighbourhood of $P(t_0)$. Thus a new current point $P'(t)$ on (P), which is coincident always with $P(t + t_0 + T')$, is within the ϵ'-neighbourhood of $P(t_1)$ at $t = 0$ and within the η'-neighbourhood of $P(t_0)$ at $t = -T'$ and $t = T''$.

Choose κ such that the distance $\overline{R(t)P'(t)}$ does not exceed η' for $-T' \leq t \leq T''$ provided only that $\overline{R(0)\ P'(0)} \leq \kappa$; and choose a positive κ' less than both κ and ϵ'. Then since $2\eta' = \eta$ any current point within the κ'-neighbourhood of $P'(0)$ i. e. of $P(t_0 + T')$ at $t = 0$ is within the η-neighbourhood of $P(t_0)$ at $t = -T', T''$. The former neighbourhood is contained by $\bar{\mathfrak{N}}$ since its radius and the distance of its centre from $P(t_1)$ are both less than ϵ', $= \frac{1}{2}\epsilon$; and it satisfies all the conditions required for $\phi(\bar{\mathfrak{N}}, \eta, T)$.

(iv) Proceeding now to the main theorem, let $\{\eta^{(i)}\}$ be a sequence of positive numbers diminishing to the limit 0, and $\{T^{(i)}\}$ an unbounded increasing sequence of positive numbers. Starting with the neighbourhood $\mathfrak{N}$ of (i), and the numbers $\eta^{(1)}, T^{(1)}$, select any closed neighbourhood $\bar{\mathfrak{N}}$ of $P(t_0)$ contained by $\mathfrak{N}$, and let a $\phi(\bar{\mathfrak{N}}, \eta^{(1)}, T^{(1)})$ be determined as in (iii). Then as in (ii) select two closed neighbourhoods $\bar{\mathfrak{N}}_0$, $\bar{\mathfrak{N}}_1$ centred at points of (P), contained by $\phi(\bar{\mathfrak{N}}, \eta^{(1)}, T^{(1)})$ and hence by $\bar{\mathfrak{N}}$, and not coupled. The radii of these may be chosen less than $\eta^{(1)}$. Then by (iii) any current point in $\bar{\mathfrak{N}}_0$ or $\bar{\mathfrak{N}}_1$ at $t = 0$ is within the $\eta^{(1)}$-neighbourhood of $P(t_0)$ at two instants at least, one prior to $t = -T^{(1)}$ and the other subsequent to $t = T^{(1)}$.

Next let there be determined a $\phi(\bar{\mathfrak{N}}_0, \eta^{(2)}, T^{(2)})$, and closed neighbourhoods $\bar{\mathfrak{N}}_{00}$, $\bar{\mathfrak{N}}_{01}$ of two points of (P), contained thereby, of radius less than $\eta^{(2)}$, and not coupled. Then $\bar{\mathfrak{N}}_{00}$, $\bar{\mathfrak{N}}_{01}$ are contained by $\bar{\mathfrak{N}}_0$, and any current point within either at $t = 0$ is within the $\eta^{(2)}$-neighbourhood of $P(t_0)$ at two instants at least, one prior to $t = -T^{(2)}$ and the other subsequent to $t = T^{(2)}$.

---

[5] This lemma is based on Birkhoff, *Dynamical Systems*, Chap. VII, sec. 11.

Next starting from $\bar{\mathfrak{N}}_1$ determine closed neighbourhoods $\bar{\mathfrak{N}}_{10}$, $\bar{\mathfrak{N}}_{11}$ therein with similar properties.

Proceeding thus, we have at the n-th stage 2^n closed neighbourhoods symbolized $\bar{\mathfrak{N}}_{i_1 i_2 \dots i_n}$, where each i is either 0 or 1. Each of these is contained in all those of preceding stages whose symbols are obtained by curtailing the suffix $i_1 i_2 \cdot \cdot \cdot i_n$, is centred at a point of (P), and is of radius less than $\eta^{(n)}$; and a current point in any one of them at $t = 0$ is within the $\eta^{(n)}$-neighbourhood of $P(t_0)$ at instants prior to $t = -T^{(n)}$ and subsequent to $t = T^{(n)}$. No two of the neighbourhoods from any stages are coupled, since this would imply the coupling of corresponding neighbourhoods with curtailed symbols, which with sufficient curtailment is contrary to the construction.

(v) Corresponding to every infinite sequence $i_1 i_2, \cdot \cdot \cdot$, where each i is 0 or 1, we have a sequence of closed neighbourhoods

$$\bar{\mathfrak{N}}, \bar{\mathfrak{N}}_{i_1}, \bar{\mathfrak{N}}_{i_1 i_2}, \bar{\mathfrak{N}}_{i_1 i_2 i_3}, \cdot \cdot \cdot$$

each contained in its predecessor. There are $2^{\aleph_0}$ distinct sequences of this sort. Since the neighbourhoods are closed, each sequence determines a point $Q_{i_1 i_2 \dots}$ common to all its members, there being only one such point since the radius $\eta^{(n)}$ tends to zero as $n \to \infty$. Each point Q belongs to $L_\omega(P)$ since it is a limit-point of the points of (P) at which the neighbourhoods of the corresponding sequence are centred. A current point which is at Q at $t = 0$ is for each n in the $\eta^{(n)}$-neighbourhood of $P(t_0)$ at instants prior to $t = -T^{(n)}$ and subsequent to $t = T^{(n)}$; hence $P(t_0)$ is both an α- and ω-limit point of (Q). Hence

$$(P) \subseteqq L_\alpha(Q), \qquad (P) \subseteqq L_\omega(Q);$$

and since the sets on the right are closed,

$$L_\omega(P) \subseteqq L_\alpha(Q), \qquad L_\omega(P) \subseteqq L_\omega(Q).$$

But since $Q \subset L_\omega(P)$,

$$L_\alpha(Q) \subseteqq L_\omega(P), \qquad L_\omega(Q) \subseteqq L_\omega(P).$$

Hence $L_\alpha(Q) = L_\omega(Q) = L_\omega(P)$, and (Q) consists of π-points.

Of the $2^{\aleph_0}$ points $Q_{i_1 i_2 \dots}$, no two can be on the same trajectory arc within $\mathfrak{N}$, for this would imply the coupling of neighbourhoods of the sequences containing the two points. Hence there are at most $\aleph_0$ of the points on any one trajectory, and the aggregate of distinct trajectories consisting of π-points has the cardinal number of the continuum.

Q. E. D.

In virtue of this theorem the stratum associated with a pervasive-asymptotic solution can be regarded as associated with a solution which is pervasive, and

necessarily *non-uniformly* pervasive (Theorem III) since the stratum contains μ- and σ-points.

THEOREM VII. (i) *An arbitrary neighbourhood of any point of the stratum $\mathfrak{S}$ associated with a non-uniformly pervasive or pervasive-asymptotic solution contains either σ- or μ-points of $\mathfrak{S}$.*

(ii) *If the σ-points are not everywhere dense in $\mathfrak{S}$ they are nowhere dense, and $\mathfrak{S}$ contains a μ-trajectory of each species: that is, a trajectory (R) for which $L_\omega(R) = \mathfrak{S}$, $L_a(R) \subset \mathfrak{S}$, and a trajectory (R') for which $L_a(R') = \mathfrak{S}$, $L_\omega(R') \subset \mathfrak{S}$.*

(i) Let $\mathfrak{N}$ be a given neighbourhood of a point of $\mathfrak{S}$. Let S be a σ-point of $\mathfrak{S}$, and $P(t)$ a current point on a trajectory of $\mathfrak{S}$ which consists of π-points. Then we can choose t_0 so that $P(t_0)$ is within $\mathfrak{N}$. $P(t_0)$ does not belong to $(S) + L_\omega(S) + L_a(S)$ since every point of this set is a σ-point. This set being closed, $P(t_0)$ has hence a positive distance from it; and there is a neighbourhood $\mathfrak{N}'$ of $P(t_0)$ contained by $\mathfrak{N}$, the distances of points of which from points of (S) have a positive lower bound.

Let $\{P(t_i)\}$ be a sequence of points of (P) converging to the point S, all external to $\mathfrak{N}'$. $P(t)$ is outside $\mathfrak{N}'$ at $t = t_i$, and is inside at some subsequent instant since $P(t_0)$ belongs to $L_\omega(P)$. Hence there is a unique t'_i corresponding to and greater than each t_i, such that $P(t)$ is outside $\mathfrak{N}'$ for $t_i \leq t < t'_i$ and is on its boundary for $t = t'_i$.

Let R be a limit point of the sequence $\{P(t'_i)\}$; then $R(t)$, the current point which is at R at $t = 0$, is for any t a limit point of the sequence $\{P(t'_i + t)\}$. Now since the distance between any points of (S) and of $\mathfrak{N}'$ has a positive lower bound, it follows from the Auxiliary Theorem that $|\, t'_i - t_i \,| \to \infty$ as the distance $\overline{SP(t_i)} \to 0$, and hence as $i \to \infty$; and with $t'_i > t_i$ this implies $t'_i - t_i \to \infty$. Hence for any fixed negative t we have, for all sufficiently large i, $t_i < t'_i + t < t'_i$; so that $P(t'_i + t)$ is outside $\mathfrak{N}'$, and hence $R(t)$ not within $\mathfrak{N}'$. Thus $L_a(R)$ has no point within $\mathfrak{N}'$, and (R) is a trajectory consisting of σ- or μ-points. The point R lies on the boundary of $\mathfrak{N}'$ since it is a limit point of the $P(t'_i)$ which lie there, and is the required σ- or μ-point within $\mathfrak{N}$.

Note: It is not proved that S belongs to $L_a(R)$, nor that $R(t)$ is for any positive t within $\mathfrak{N}'$.

(ii) Suppose now that the σ-points are not everywhere dense in $\mathfrak{S}$, and that $\mathfrak{N}$ is a neighbourhood of a point of $\mathfrak{S}$ free from σ-points. Let $\mathfrak{D}$ be any neighbourhood of any point of $\mathfrak{S}$.

We can choose t_0, t_1 so that $P(t_0)$, $P(t_1)$ lie within $\mathfrak{N}$, $\mathfrak{D}$ respectively.

A sufficiently small neighbourhood $\mathcal{D}'$ of $P(t_1)$ has its $(t_0 - t_1)$-transform contained by $\mathcal{N}$, and $\mathcal{D}'$ contains no σ-points since any transform of a σ-point is a σ-point. Hence there is a sub-neighbourhood of $\mathcal{D}$, centred at a point of $\mathcal{S}$, free of σ-points; and the σ-points are nowhere dense in $\mathcal{S}$.

It follows that the point R defined in (i), being within $\mathcal{N}$, must be a μ-point. Hence (R) consists of μ-points, and since $L_\alpha(R) \subset \mathcal{S}$ we must have $L_\omega(R) = \mathcal{S}$. In a similar manner, by taking t'_i less than t_i, we find a trajectory (R') with $L_\alpha(R') = \mathcal{S}$, $L_\omega(R') \subset \mathcal{S}$.

Q. E. D.

In view of this theorem there are three logically possible types of strata $\mathcal{S}$ associated with non-uniformly pervasive solutions, the characteristic properties being

(a) the σ-points are nowhere dense and the μ-points everywhere dense in $\mathcal{S}$,

(b) both σ- and μ-points are everywhere dense in $\mathcal{S}$,

(c) the σ-points are everywhere dense and there are no μ-points in $\mathcal{S}$.

If the solution with which $\mathcal{S}$ was originally associated is pervasive-asymptotic, $\mathcal{S}$ can be of type (a) or (b) only. The type (b) can be subdivided according as there are μ-trajectories of both species or of one species only.

In §§ 9-11 below, all these types are illustrated, though only in the cases of type (a) are the X_i in the differential equations regular-analytic. The type (b) is illustrated by the quasi-ergodic systems of geodesics on closed surfaces of negative curvature which have been given by Artin [1] and others; in such cases the stratum is not of course the surface on which the geodesics lie, but a related three-dimensional manifold.

(iii) In the case where the σ-points are everywhere dense in $\mathcal{S}$, let us call a point Q of $\mathcal{S}$ a *σ-point of class η* if η is the upper bound of distances of points of $\mathcal{S}$ from the trajectory (Q); and let Σ_η be the set of all σ-points of $\mathcal{S}$ whose class is not less than η. Then it is easy to establish the following facts:

(a) *If Q belongs to Σ_η then every point of (Q), $L_\omega(Q)$ and $L_\alpha(Q)$ belongs to Σ_η.*

(b) *There is no point of Σ_η in a sufficiently small neighbourhood of any π-point or μ-point or σ-point of class less than η.*

(c) *The set Σ_η is closed.*

(d) *Either an arbitrary neighbourhood of Σ_η contains complete σ-trajectories (whose class-number is less than but arbitrarily near η), or there are trajectories of positively (negatively) semi-asymptotic type whose ω- (α-)limit set belongs to Σ_η; in the latter alternative a set of suitably selected points, one on each*

of the semi-asymptotic trajectories, is closed, but the set of all points on these trajectories is not necessarily closed.

Birkhoff has called the system of trajectories filling a closed connected region $\mathfrak{d}$ *transitive* if arbitrary neighbourhoods of any two points of the region are traversed by the same trajectory of the system; and he has proved (*Dynamical Systems*, Chap. VII, Sec. 11) that in this case $\mathfrak{d}$ contains trajectories whose ω- and α-limit sets coincide with $\mathfrak{d}$. Thus $\mathfrak{d}$ is a stratum, and the theories of the structure of strata and of transitive systems are identical. In Birkhoff's terminology, the trajectories which consist of π-points are *general*, and those which consist of σ- or μ-points are *special*. For the non-trivial case in which the system contains special trajectories, it may be observed that Theorem VI contains a new result concerning the power of the aggregate of general trajectories, and that Theorem VII reproduces Birkhoff's theorems on the special trajectories (*Dynamical Systems*, Chap. VII, Sec. 9) with a somewhat more direct proof.

The analysis of $\mathfrak{d}$ into π-, μ- and σ-points may be of some significance in connection with the problem of metrical transitivity. If $\mathfrak{d}$ is a regular hypersurface on which measure is suitably defined and on which there is a positive integral-invariant, it is known (Carathéodory [1]) that the μ-points are of measure 0. The question then arises whether the σ-points are of measure 0, which is of course necessary for the metrical transitivity of $\mathfrak{d}$. I have not been able to do more than reduce this question, by means of (c) in (iii) above, to the following: Is a *closed* invariant set which is a part only of $\mathfrak{d}$ necessarily of measure 0?

Construction of Systems of Differential Equations Having Non-Uniformly Pervasive and Pervasive-Asymptotic Solutions.

---

8. As starting-point of the construction we take a system of differential equations in two dependent variables r, z:

$$dr/ds = R(r, z), \qquad dz/ds = Z(r, z) \tag{6}$$

which possesses a (real) periodic solution

$$r = f(s), \qquad z = g(s). \tag{7}$$

Such systems are easy to construct, by starting with an equation

$$F(r, z) = \text{constant} \tag{8}$$

which, with r and z as rectangular coördinates, represents a family of closed

curves without singular points, and eliminating the constant by differentiation. For example, from

$$r^2 + z^2 = \text{constant} \tag{9}$$

we may pass to the differential equations

$$dr/ds = -z, \qquad dz/ds = r$$

whose solutions

$$r = c\cos s, \qquad z = c\sin s$$

are periodic, and reproduce the circles (9).

The properties postulated for the functions occurring in (6) and (7) are as follows:

(i) $R(r, z)$ and $Z(r, z)$ are analytic functions of r, z, real and regular at all points $(f(s), g(s))$ for which s is real.

(ii) $f(s)$ and $g(s)$ are analytic functions of s, real and regular for all real s, and periodic with period S;

(iii) the lower bounds of $f(s)$ and $g(s)$, for s real, are positive. The properties (i), (ii) are here stated in a form convenient for subsequent requirements, and without regard for the fact that they have some logical connection. From them, together with the general theorem of uniqueness of solution of a system such as (6), there follows:

(iv) the closed curve specified by (7) has no singular points.

The property (iii) is, in geometrical terms, that the curve (7) lies entirely in the positive quadrant of the rz-plane.

We now pass to a system with three dependent variables x, y, z by adjoining to (6) a differential equation

$$d\theta/ds = 2\pi/T$$

where T is any constant such that S/T is an irrational number, and putting

$$x = r\cos\theta, \qquad y = r\sin\theta. \tag{10}$$

In the space in which x, y, z are rectangular coördinates, r, θ, z are cylindrical coördinates, and the equations (7) represent a torus $\mathfrak{S}$ obtained by rigid rotation of the curve (7) about the axis $r = 0$; and in virtue of (iii), (iv) above, $\mathfrak{S}$ has no singular points.

In terms of x, y, z the system is

$$dx/ds = xR/r - 2\pi y/T, \quad dy/ds = yR/r + 2\pi x/T, \quad dz/ds = Z, \tag{11}$$

where we are to put $r = +\sqrt{(x^2 + y^2)}$; its right members are (from (i) and

(iii)) single-valued regular functions of x, y, z for (x, y, z) on $\mathfrak{S}$. $\mathfrak{S}$ is made up of a family of the trajectories of (11), which may be specified by the corresponding solutions

$$(12) \quad x = f(s)\cos(2\pi s/T + \theta_0), \quad y = f(s)\sin(2\pi s/T + \theta_0), \quad z = g(s),$$

θ_0 being an arbitrary constant.

It is well known that, since S/T is irrational, this family is a minimal set of recurrent, or uniformly pervasive, trajectories each of which has $\mathfrak{S}$ for associated stratum. Each trajectory, if followed in either sense, spirals indefinitely round $\mathfrak{S}$ without singularity, and has infinitely many arcs lying in an arbitrary neighbourhood of any point of $\mathfrak{S}$. Each trajectory may be specified by any value of the constant θ_0 chosen from an enumerably infinite set; and there are an infinity of distinct trajectories, of the cardinal number of the continuum, one of which passes through any assigned point of $\mathfrak{S}$.

9. Choose any point A of the surface $\mathfrak{S}$, say (x_1, y_1, z_1). Let $E(x, y, z)$ be an analytic function, real and regular for all (x, y, z) on $\mathfrak{S}$ and positive at all these points except A, where it vanishes; for example

$$E \equiv (x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2.$$

We shall shew that *the differential equations*

$$(13) \quad dx/dt = E(xR/r - 2\pi y/T), \quad dy/dt = E(yR/r + 2\pi x/T), \quad dz/dt = EZ$$

have certain solutions, viz. those which lie in $\mathfrak{S}$, which are either non-uniformly pervasive or pervasive-asymptotic.

Since (13) is related to (11) by the substitution

$$dt = ds/E,$$

the trajectories of (13) are geometrically coincident with those of (11); and for the solutions carried by these trajectories we have to determine the manner in which t is correlated with their points. For the solutions in $\mathfrak{S}$ we have

$$(14) \quad t = \int_{s_0}^{s} \frac{ds}{E\{f(s)\cos(2\pi s/T + \theta_0), f(s)\sin(2\pi s/T + \theta_0), g(s)\}},$$

where s_0 is an arbitrary constant effectively additive with t.

First let θ_0 have such a value that the curve (12) passes through A, the value, necessarily unique, of the parameter s at this point being s_1 say. Then for $s = s_1$ the denominator E on the right of (14) vanishes, while for $s \neq s_1$ it is positive with a positive upper bound. Since E is a regular function of x, y, z which are, by (12), regular functions of s, E is a regular function of s; and near s_1 we have a convergent expansion

$$E = a(s - s_1)^2 + b(s - s_1)^3 + \cdots$$

where, because E is never negative on $\mathcal{S}$, the leading term has a positive coefficient and an even exponent. Hence the integral in (14) diverges at $s = s_1$ like an odd negative-integral power of $(s_1 - s)$.

Thus if $s_0 > s_1$, (14) gives t as a single-valued strictly increasing function of s defined for $s > s_1$, and $t \to -\infty$ as $s \to s_1 + 0$; while since $1/E$ has a positive lower bound, $t \to \infty$ as $s \to \infty$. If $s_0 < s_1$, t is a single-valued strictly increasing function of s defined for $s < s_1$, and $t \to \infty$ as $s \to s_1 - 0$, $t \to -\infty$ as $s \to -\infty$. In both cases, inversion gives s as a single-valued strictly increasing function of t defined for all real t; if $s_0 > s_1$ we have $s \to s_1$ as $t \to -\infty$, and $s \to \infty$ as $t \to \infty$; if $s_0 < s_1$ we have $s \to -\infty$ as $t \to -\infty$, and $s \to s_1$ as $t \to \infty$. Hence the curve (12) splits into three complete trajectories of (13), for which respectively $s = s_1$ (giving the equilibrium point A), $s > s_1$, and $s < s_1$. Any solution carried by the portion for which $s > s_1$ has the point A as its α-limit set and, since $s \to \infty$ as $t \to \infty$, the whole surface $\mathcal{S}$ as its ω-limit set. For the other portion the ω-limit set is A and the α-limit set is $\mathcal{S}$. Thus for (13) we have just one pervasive-asymptotic trajectory of each species, each of which carries of course the infinitely many trivially different solutions corresponding to different values of s_0.

Now let θ_0 have such a value that the curve (12) passes arbitrarily near to but not through A. Then in (14), E vanishes for no value of s though its lower bound is 0, and t is a single-valued strictly increasing function of s defined for all real s; and since $1/E$ has a positive lower bound, $t \to \pm\infty$ as $s \to \pm\infty$. Inversion gives s as a single-valued strictly increasing function of t, defined for all real t, with $s \to \pm\infty$ as $t \to \pm\infty$. Thus the curve furnishes one trajectory of (13), any solution carried by which has the whole surface $\mathcal{S}$ as its α- and as its ω-limit set. The solution (or trajectory) is hence pervasive, with $\mathcal{S}$ as its stratum. That it is non-uniformly pervasive follows most easily from Theorem III (i) since we already know $\mathcal{S}$ to contain σ- and μ-points. The mechanism of the non-uniformity is of course that in (14) the denominator E is arbitrarily small for the infinitely many values of s corresponding to sufficiently near approaches to A; and it may be proved directly that if a current point enters the ϵ-neighbourhood of A, where $\epsilon \sim 0$, it remains within a fixed neighbourhood of A for a time of order $1/\epsilon$.

We see therefore that $\mathcal{S}$ is the stratum associated with each of the non-uniformly *pervasive* or *pervasive*-asymptotic trajectories of (13) contained thereby. It contains one σ-point A, two trajectories made up of μ-points, and an infinity, of the cardinal number of the continuum, of trajectories made up of π-points.

10. The example of § 9 is in accord with the definitions of §§ 3-5, but is, so to speak, psychologically weak: the equilibrium point A is removable by change of the independent variable, and it might be held that the definitions should be modified so as to presuppose such removal. It is however easy to pass to an example free from this possible objection.

Adjoin to (13) the differential equation

$$d\psi/dt = 2\pi/U,$$

where U is a positive constant, and let new variables u, v be introduced by the relations

$$(15) \qquad u = z\cos\psi, \quad v = z\sin\psi, \quad z = +\sqrt{(u^2+v^2)}.$$

We obtain the system of the fourth order

$$(16) \qquad \left.\begin{array}{ll} dx/dt = E(xR/r - 2\pi y/T), & dy/dt = E(yR/r + 2\pi x/T) \\ du/dt = uEZ/z - 2\pi v/U, & dv/dt = vEZ/z + 2\pi u/U \end{array}\right\},$$

of which ∞^1 solutions are got by adjoining to any solution of (13) the equation

$$\psi = 2\pi t/U + \psi_0.$$

If the solution of (13) lies in $\mathfrak{s}$, the solutions of (16) thus generated lie in the hypersurface $\mathfrak{H}$ obtained by rotating $\mathfrak{s}$ about the plane $u = v = 0$ (or $z = 0$) in the four dimensional space in which x, y, u, v, are rectangular coördinates. $\mathfrak{H}$ is without singular points since by the initial assumption (iii) of § 8 $\mathfrak{s}$ has no point common with the plane $z = 0$; and $\mathfrak{s}$ is the section of $\mathfrak{H}$ by the half-hyperplane $\psi = 0$. Any solution (Q) in $\mathfrak{H}$ has in common with $\mathfrak{s}$ the points

$$\cdots P(t_0 - 2U),\ P(t_0 - U),\ P(t_0),\ P(t_0 + U),\ P(t_0 + 2U), \cdots$$

occupied at equidistant time-intervals by some current point $P(t)$ which moves in $\mathfrak{s}$ according to the differential equations (13). The points of $L_\omega(Q)$, $L_\alpha(Q)$ in $\mathfrak{s}$ are the limit points of the respective sequences

$$(17) \qquad P(t_0), P(t_0 + U), P(t_0 + 2U), \cdots,$$

$$(18) \qquad P(t_0), P(t_0 - U), P(t_0 - 2U), \cdots,$$

and the complete sets $L_\omega(Q)$, $L_\alpha(Q)$ consist of the trajectories of (16) through these limit points. We call (17) and (18) the *positive* and *negative U-sets based on* $P(t_0)$ and desire therefore to know their derived sets when $P(t_0)$ is any assigned point of $\mathfrak{s}$.

If $P(t_0)$ is at A all the points of (17), (18) are at A, and we have for the system (16) a periodic trajectory $\mathfrak{J}$ carrying the ∞^1 trivially-different solutions

$$(19) \qquad x = x_1,\ y = y_1,\ u = z_1\cos(2\pi t/U + \psi_0),\ v = z_1\sin(2\pi t/U + \psi_0).$$

If $P(t)$ asymptotes to A as $t \to \infty$, any associated positive U-set con-

verges to A, and the generated solutions of (16) have the periodic trajectory $\mathfrak{J}$ for ω-limit set. There are ∞^2 such solutions, carried by ∞^1 trajectories. There are similarly ∞^1 trajectories of (16) having $\mathfrak{J}$ as α-limit set.

The only other, and most general, case is that where $P(t)$ passes arbitrarily near to each point of $\mathfrak{S}$ as $t \to \infty$ (or as $t \to -\infty$). In this case it is clear that *any associated positive (or negative) U-set cannot converge to A*, but it is not immediately evident whether or not such a set is everywhere dense in $\mathfrak{S}$. [If for instance, altering our hypotheses, we take $E \equiv 1$ on $\mathfrak{S}$, such a set is not or is everywhere dense in $\mathfrak{S}$ according as U is or is not commensurably related to S and T.] It can be proved that *every U-set except those which converge to A is everywhere dense in $\mathfrak{S}$, whatever the value of the constant U, provided that S/T is a transcendental number of a certain type and that the analytic function E of the complex variable s is suitably chosen.*[6] The proof of this is too long to be given here, and the following result is sufficient for our purpose:

THEOREM VIII. *For any choice, in accordance with the conditions previously stipulated, for the functions and constants occurring in the right members of the equations* (13), *there is a value of U arbitrarily near to any given positive number U_0 for which every U-set in $\mathfrak{S}$, except those which converge to the point A, is everywhere dense in $\mathfrak{S}$.*

Assuming this for the moment, it follows that any positively (negatively) semi-pervasive solution of (13) generates ∞^1 solutions of (16) which are positively (negatively) semi-pervasive with $\mathfrak{H}$ as the associated stratum. Thus $\mathfrak{H}$ is made up of the following sets: (a) the periodic trajectory $\mathfrak{J}$; (b) ∞^1 trajectories (Q) for which $L_\omega(Q) = \mathfrak{J}$, $L_\alpha(Q) = \mathfrak{H}$, and ∞^1 trajectories (Q') for which $L_\alpha(Q') = \mathfrak{J}$, $L_\omega(Q') = \mathfrak{H}$; and (c) ∞^2 trajectories (Q'') for which $L_\omega(Q'') = L_\alpha(Q'') = \mathfrak{H}$. It follows as for the system (13) that the trajectories (Q'') are non-uniformly pervasive.

Proof of Theorem VIII. (i) If $P(t)$ is any current point in $\mathfrak{S}$, we call the aggregate of points $\{P(t)\}$ for which $t' \leq t \leq t' + \tau$ a *closed*[7] *trajectory-arc of duration τ*. If Q is any point of $\mathfrak{S}$, other than A, let $\mathfrak{C}$ be an analytic arc through Q lying in $\mathfrak{S}$ and nowhere tangential to a trajectory; the U-transform of $\mathfrak{C}$ is then an analytic arc $\mathfrak{C}_U$, and the closed trajectory-arcs which stretch from $\mathfrak{C}$ to $\mathfrak{C}_U$ are all of duration U.

---

[6] If $S/T = \dfrac{1}{a_1 +} \dfrac{1}{a_2 +} \dfrac{1}{a_3 +} \cdots$ with $a_{n+1} > 2q_n/\epsilon_{n+1}$, where p_n/q_n is the n-th convergent and $\epsilon_n \to 0$ as $n \to \infty$, asymptotic formulae connecting the values of the integral in (14) over suitably chosen ranges can be developed by deformation of the path of integration, and lead to the stated result.

[7] i. e. "closed" in the sense that the end points belong to the arc.

If $P(t)$ passes arbitrarily near to each point of $\mathfrak{S}$ as $t \to \infty$, the positive U-set based on any point $P(t_0)$ of its trajectory has a limit point on any closed trajectory-arc QQ_U of duration U. For the half-trajectory described by $P(t)$ for $t \geq t_0$ cuts $\mathscr{C}$, at $P(t')$ say, in an arbitrary neighbourhood of Q. $P(t'+U)$ then lies on $\mathscr{C}_U$, and if $P(t')$ is sufficiently near to Q the closed trajectory-arc from $P(t')$ to $P(t'+U)$ lies in an arbitrary neighbourhood of the arc QQ_U. The first named arc contains one point of the U-set

$$P(t_0), P(t_0+U), P(t_0+2U), \cdots$$

which is in question, or two points if $t' \equiv t_0 \pmod{U}$. Hence in an arbitrary neighbourhood of the closed arc QQ_U there are infinitely many points of the U-set, which has therefore a limit point on the arc.

Similarly, *if $P(t)$ passes arbitrarily near to each point of $\mathfrak{S}$ as $t \to -\infty$, the negative U-set based on any point $P(t_0)$ of its trajectory has a limit point on any closed trajectory-arc of duration U.*

(ii) *If $P(t)$ passes arbitrarily near to each point of $\mathfrak{S}$ as $t \to \infty$, $P(t_0)$ is a given point of its trajectory, and U_0 is a given positive number, then we can choose U as near as we please to U_0 such that the positive U-set based on $P(t_0)$ has every point of the trajectory-arc QQ_U as a limit point.*

We shew how to choose U so that an arbitrary neighbourhood of any point of the trajectory-arc from Q to its $(2U_0)$-transform Q_{2U_0} contains infinitely many points of the positive U-set based on $P(t_0)$. For this purpose let $\mathscr{C}$ be an analytic arc through Q not tangential to any trajectory, ϕ the arc-length on $\mathscr{C}$ measured from Q, and $\mathscr{C}_r$ the r-transform of $\mathscr{C}$. Any trajectory-arc near QQ_{2U_0} can then be referred to the value of ϕ at the point at which it cuts $\mathscr{C}$. A set of domains $\mathfrak{D}_n$ in $\mathfrak{S}$ is then defined as follows, η being a positive constant:

$\mathfrak{D}_1$	bounded	by	arcs	of	$\mathscr{C}$,	$\mathscr{C}_{U_0}$	and	of	trajectories	$\phi = \eta, -\eta$
$\mathfrak{D}_2$	"	"	"	"	$\mathscr{C}_{U_0}$,	$\mathscr{C}_{2U_0}$	"	"	"	$\phi = \pm \frac{1}{2}\eta$
$\mathfrak{D}_3$	"	"	"	"	$\mathscr{C}$,	$\mathscr{C}_{\frac{1}{2}U_0}$	"	"	"	$\phi = \pm \frac{1}{3}\eta$
$\mathfrak{D}_4$	"	"	"	"	$\mathscr{C}_{\frac{1}{2}U_0}$,	$\mathscr{C}_{U_0}$	"	"	"	$\phi = \pm \frac{1}{4}\eta$
$\mathfrak{D}_5$	"	"	"	"	$\mathscr{C}_{U_0}$,	$\mathscr{C}_{\frac{3}{2}U_0}$	"	"	"	$\phi = \pm \frac{1}{5}\eta$
$\mathfrak{D}_6$	"	"	"	"	$\mathscr{C}_{\frac{3}{2}U_0}$,	$\mathscr{C}_{2U_0}$	"	"	"	$\phi = \pm \frac{1}{6}\eta$
$\mathfrak{D}_7$	"	"	"	"	$\mathscr{C}$,	$\mathscr{C}_{\frac{1}{4}U_0}$	"	"	"	$\phi = \pm \frac{1}{7}\eta$.
.	.	.	.	.	.	.	.	.	.	.

An arbitrary neighbourhood of any point of the arc QQ_{2U_0} then contains infinitely many of the $\mathfrak{D}_n$.

Let $2t_n$ be the time-interval between the end-arcs of $\mathfrak{D}_n$, so that

$$t_1 = t_2 = \tfrac{1}{2}U_0,\ t_3 = \cdots = t_6 = \tfrac{1}{4}U_0,\ t_7 = \cdots = t_{14} = \tfrac{1}{8}U_0, \text{ etc.};$$

and

$$t_{n+1} \leq t_n < 2U_0. \tag{20}$$

Call the arc in $\mathfrak{D}_n$ which is at equal time-intervals from its end-arcs its *axis*; e. g. $\mathfrak{C}_{1/2U_0}$ is the axis of $\mathfrak{D}_1$.

By hypothesis $P(t)$ traverses each $\mathfrak{D}_n$ for arbitrarily large values of t, so there are arbitrarily large integers i_n such that the point $P(t_0 + i_n U)$ of the U-set based on $P(t_0)$ lies on a trajectory-arc traversing $\mathfrak{D}_n$ at a time-interval less than U from its axis. If U increases by δU, t_0 remaining fixed, $P(t_0 + i_n U)$ advances along the trajectory (P) through an arc of duration $i_n \delta U$; so we can choose δU, with $|\delta U| < U/i_n$, such that $P(t_0 + i_n(U + \delta U))$ lies in $\mathfrak{D}_n$ on its axis. By giving U a succession of increments from the starting-value U_0 we thus move points of the U-set, of suitably chosen ranks $i_1, i_2, \cdots$, in succession on to the axes of $\mathfrak{D}_1, \mathfrak{D}_2, \cdots$. The alteration in U at the n-th stage produces a consequential displacement of each of the points $P(t_0 + i_1 U), \cdots, P(t_0 + i_{n-1} U)$; but by taking the ratios i_2/i_1, $i_3/i_2, \cdots$ sufficiently large we secure that, for each j, $P(t_0 + i_j U)$ never gets moved out of $\mathfrak{D}_j$.

A suitable rule for choosing $i_1, i_2, \cdots$ is as follows: Choose a constant κ between 0 and $\frac{1}{2}$. Choose i_1 greater than $2/\kappa$ and such that $P(t_0 + i_1 U_0)$ lies on a trajectory-arc traversing $\mathfrak{D}_1$, at a time-interval less than U_0 from its axis. Choose $U_1, = U_0 + \delta_0 U$, such that $P(t_0 + i_1 U_1)$ lies on the axis of $\mathfrak{D}_1$. Then

$$|\delta_0 U| < U_0/i_1 < 2U_0/i_1 < \kappa U_0 < U_0, \qquad 0 < U_1 < 2U_0.$$

Suppose that $i_1, \cdots, i_{n-1}$ have been chosen in succession such that

(a) $U_{n-1} = U_0 + \delta_0 U + \delta_1 U + \cdots + \delta_{n-2} U$ with
$|\delta_j U| < \kappa^{j+1} U_0, \qquad (j = 0, 1, \cdots, n-2),$

so that

$$|U_{n-1} - U_0| < \kappa U_0/(1-\kappa) < U_0; \tag{21}$$

(22) (b) $i_j > 2U_0 i_{j-1}/\kappa t_{j-1}, \qquad (j = 2, 3, \cdots, n-1)$;

(c) $P(t_0 + i_{n-1} U_{n-1})$ lies on the axis of $\mathfrak{D}_{n-1}$;

(d) $P(t_0 + i_j U_{n-1})$, for $j = 1, 2, \cdots, n-2$, lies in $\mathfrak{D}_j$ at a time-interval from its axis less in absolute value than

$$(\kappa + \kappa^2 + \cdots + \kappa^{n-1-j}) t_j.$$

Choose i_n greater than $2U_0 i_{n-1}/(\kappa t_{n-1})$ and such that $P(t_0 + i_n U_{n-1})$ lies on a trajectory-arc traversing $\mathfrak{D}_n$, at a time-interval from its axis less than U_{n-1}. Choose $U_n, = U_{n-1} + \delta_{n-1} U$, such that $P(t_0 + i_n U_n)$ lies on the axis of $\mathfrak{D}_n$. Then from (21)

$$|\delta_{n-1} U| < U_{n-1}/i_n < 2U_0/i_n;$$

and the time-interval between $P(t_0 + i_j U_{n-1})$ and $P(t_0 + i_j U_n)$ is $i_j \delta_{n-1} U$, with an absolute value less than $2U_0 i_j / i_n$. From (20), (22) we have, for $j = 2, 3, \cdots, (n-1)$,

$$i_{j-1}/i_j < \kappa(t_{j-1}/2U_0) < \kappa,$$

and the same relation holds for $j = n$ in virtue of the choice of i_n. Hence

$$2U_0/i_n = 2U_0/i_1 \cdot i_1/i_2 \cdot \cdots i_{n-1}/i_n < \kappa U_0 \cdot \kappa^{n-1} = \kappa^n U_0,$$

and for $j = 1, 2, \cdots, n-1$

$$\begin{aligned} i_j/i_n &= i_j/i_{j+1} \cdot i_{j+1}/i_{j+2} \cdot \cdots i_{n-1}/i_n \\ &< (\kappa t_j/2U_0) \cdot \kappa^{n-j-1} = \kappa^{n-j} t_j/2U_0. \end{aligned}$$

Thus $| \delta_{n-1} U | < \kappa^n U_0$, and the time-interval of $P(t_0 + i_j U_n)$ from the axis of $\mathfrak{D}_j$ is less in absolute value than $(\kappa + \cdots + \kappa^{n-j}) t_j$. The properties (a), (b), (c), (d) are thus by induction established for all n. The sequence $\{U_n\}$ converges to the desired value U, for which

$$| U - U_0 | < \kappa U_0/(1-\kappa) < U_0, \qquad 0 < U < 2U_0,$$

while the time-interval of $P(t_0 + i_j U)$ from the axis of $\mathfrak{D}_j$ is less in absolute value than $\kappa t_j/(1-\kappa)$, which is less than t_j, so that $P(t_0 + i_j U)$ lies in $\mathfrak{D}_j$. For a sufficiently small choice of κ, U is as near as we please to U_0.

It follows immediately that *every point of the trajectory-arc QQ_{2U_0} is a limit point of the U-set based on $P(t_0)$*; and there is the further relevant fact that *this arc contains a closed arc QQ_U of duration U.*

(iii) *With the value of U just chosen, every U-set is everywhere dense in $\mathfrak{A}$, except the positive (negative) U-sets on the trajectory positively (negatively) asymptotic to A.*

If any U-set has a limit point R, the Auxiliary Theorem shews that the t-transform of the set has the t-transform of R as a limit point, for any t. In particular the U-transform and the $(-U)$-transform of R are limit points *of the original set.*

Since the U-set $P(t_0), P(t_0 + U), \cdots$ which is in question in (ii) above has every point of the trajectory-arc QQ_U as a limit point, and Q_U is the U-transform of Q, it follows that every point of (Q) is a limit point. But $Q(t)$ passes arbitrarily near to each point of $\mathfrak{A}$ as $t \to \infty$, or $t \to -\infty$, or both, since every solution (except A) in $\mathfrak{A}$ has $\mathfrak{A}$ as its ω- or α-limit set, or both. Hence there is one U-set $P(t_0), P(t_0 + U), \cdots$ whose derived set is $\mathfrak{A}$.

From what was first pointed out it follows that for any t' the positive U-set based on $P(t_0 + t')$ has for derived set the t'-transform of $\mathfrak{A}$, which coincides with $\mathfrak{A}$.

Now any U-set K which is not of the type excepted in the enunciation

has by (i) a limit point, $P(t_0 + t')$ say, on (P). Hence as before K has each of the points

$$P(t_0 + t'),\ P(t_0 + t' + U),\ P(t_0 + t' + 2U), \cdots$$

as a limit point. These points having just been seen to be everywhere dense in $\mathfrak{S}$, K must be everywhere dense in $\mathfrak{S}$.

Q. E. D.

It is of interest to note that the functions R, Z, E can be so chosen, in accordance with the conditions of §§ 8, 9, that the right members of the equations (13) are polynomials in x, y, z, and those of the equations (16) polynomials in x, y, u, v. The simplest choice of $F(r, z)$ appears to be

$$F(r, z) = r^2 + z^2 + 1/r^2 + 1/z^2.$$

If $c > 4$ the curve $F = c$ consists of ovals symmetrically placed one in each quadrant; for $c = 4$ we get the four acnodes $r = \pm 1$, $z = \pm 1$; and the axes $r = 0$, $z = 0$ correspond to $c = \infty$. Eliminating c, the system (6) is

$$dr/ds = r^3(z^4 - 1) = R, \qquad dz/ds = z^3(1 - r^4) = Z.$$

The solutions for $c > 4$ are periodic with a period S which is an analytic function of c, and any one of these solutions can serve as the basis (7) for the construction. The right members of (13), (16) will be polynomials as stated if E is a polynomial in x, y, z^2; for example

$$E = (x - x_1)^2 + (y - y_1)^2 + (z^2 - z_1^2)^2.$$

11. The principles exemplified in §§ 9, 10 can be used to construct more complicated examples. The function E can be chosen, as a polynomial, to vanish at any finite number of assigned points $A_1, \cdots, A_n$ of $\mathfrak{S}$ and be positive elsewhere on $\mathfrak{S}$, and the points $A_1, \cdots, A_n$ may be on the same or on different trajectories of the system (11). If $A_1, \cdots, A_m$ are on the same trajectory, this splits into $(2m + 1)$ complete trajectories of the system (13), namely the equilibrium points $A_1, \cdots, A_m$, the finite open arcs $A_1A_2, \cdots, A_{m-1}A_m$, and two semi-infinite arcs. The points of the latter become μ-points, and $A_1, \cdots, A_m$ and the points of the finite arcs become σ-points.

If the points A are infinite in number but not everywhere dense in $\mathfrak{S}$, an E which is not identically zero must cease to be *regular-analytic* at their limit points, but can be chosen so as to be *regular*, in the sense defined in § 1, at all points. With this sort of E we can exemplify the types (b), (c) of § 7; the examples already given are of type (a), with the σ-points nowhere dense in $\mathfrak{S}$.

Taking the trajectory (12) of the system (11) for which θ_0 has an assigned value, let equilibrium points $A_1, A_2, \cdots$ be introduced at the points

$$s = s_1, s_2, \cdots, (s_1 < s_2 < s_3 \cdots),$$

where s_n is so chosen that the sequence $\{A_n\}$ has A_1 as its sole limit point; this implies that $s_n \to \infty$ as $n \to \infty$. E is of course to vanish at $A_1, A_2, \cdots$ and be positive elsewhere on $\mathfrak{S}$. Then the half-trajectory $s \geq s_1$ of (11) furnishes an infinite number of complete trajectories of (13), viz. the points A_n and the finite open arcs A_nA_{n+1}, and each of these consists of σ-points. The half-trajectory $s < s_1$ consists as before of μ-points. Hence the stratum $\mathfrak{S}$ is of type (b), with both σ- and μ-points everywhere dense therein, and all μ-points lie on one positively semi-asymptotic and negatively semi-pervasive trajectory.

If we now introduce another equilibrium point A_0 on a different trajectory of (11), this gives for (13) another σ-point A_0 and two pervasive-asymptotic trajectories, one of each species. Hence $\mathfrak{S}$ has now the σ-points everywhere dense and μ-trajectories of both species.

If instead of introducing A_0 in this manner we treat a complete trajectory (instead of a half-trajectory) of (11) in the previous manner, this is split into an infinite number of equilibrium points A_n and finite open arcs A_nA_{n+1} $(n = \cdots -2, -1, 0, 1, 2, \cdots)$, where the A_n can be so chosen that the aggregate $\{A_n\}$ has A_0 as sole limit point. Choosing E so as to be positive on $\mathfrak{S}$ except at the A_n, the stratum $\mathfrak{S}$ is of type (c), with everywhere dense σ-points and no μ-points.

MELBOURNE UNIVERSITY,
AUSTRALIA.

REFERENCES.

E. Artin [1], "Ein Mechanisches System mit quasiergodischen Bahnen," *Abh. Math. Sem. Hamburg*, bd. 3 (1923), p. 170.

G. D. Birkhoff [1], "Quelques théorèmes sur le mouvement des systèmes dynamiques," *Bull. Soc. Math. France*, t. 40 (1912), p. 305.

[2], *Dynamical Systems*, New York (1927).

G. D. Birkhoff and P. A. Smith [1], "Structure analysis of surface transformations," *Jour. de Math. Pures et Appl.*, [9], t. 7 (1928), p. 345.

C. Carathéodory [1], "Über den Wiederkehrsatz von Poincaré," *Sitz. preuss. Akad. de Wiss.*, Berlin (1919), p. 580.

J. Hadamard [1], "Sur certaines propriétés des trajectoires en Dynamique," *Jour. de Math. Pures et Appl.*, [5], t. 3 (1897), p. 331.

H. Hilmy [1], "Sur les ensembles quasi-minimaux dans les systèmes dynamiques," *Annals of Math.*, vol. 37 (1936), p. 899.

E. Husson [1], *Les trajectoires de la dynamique*, Paris (1932).

H. Poincaré [1], "Mémoire sur les Courbes définies par une Équation différentielle," *Jour. de Math. Pures et Appl.*, [3], t. 8 (1882), p. 251.

[2], *Les Méthodes Nouvelles de la Mécanique Céleste*, t. 3 (Paris, 1899), Chap. 26.

THEORY OF QUASI-GROUPS.*

By B. A. Hausmann and Oystein Ore.

The present paper contains an investigation of the theory of groupoids, i. e. algebraic systems in which one operation is defined. There exists already a number of publications on such systems and various interesting although disconnected and scattered results have been obtained. In most of these studies certain axiomatic rules have been laid down and the properties of the corresponding systems have been deduced.

Here we shall try to obtain a more systematic theory by a different approach to the problem. Instead of starting from a chosen set of axioms we analyse the conditions necessary in order to obtain specific theories and theorems. In the end this turns out to be mainly an analysis of the associative law and the various possible formulations of the associative law represent the conditions for the existence of various analogues to the theorems in ordinary group theory.

In the first chapter one finds a classification of the groupoids and the definition of quasi-groups. Then follows an analysis of the conditions for co-set expansions with various properties, and one is led to three different types of associative laws. An interesting phenomenon by such quasi-groups is the existence of a minimal quasi-group not equal to a unit element but contained in all other quasi-groups.

In the second chapter one finds the structure theory of quasi-groups. We define normal quasi-groups and consider the conditions for laws of isomorphism. Next we determine the conditions for the normal quasi-groups to form a Dedekind structure and from these investigations follows from the general structure theory the conditions for the theorem of Jordan-Hölder and the direct decomposition theorem of Schmidt-Remak. Conditions for the existence of quotient groupoids are also given. Various problems which we had to omit from these considerations have been mentioned in the concluding section.

Chapter I. Quasi-Groups and the Associative Law.

1. Definitions and axioms. Since there exists a great variety of nomenclatures in our subject it may be well to specify at the beginning the definitions of the fundamental concepts used in the following.

* Received May 23, 1937.

We consider systems $\mathfrak{A}$ consisting of an arbitrary finite or infinite number of elements $a_1, a_2, \cdots$. Such a system shall be called a *groupoid* if it satisfies the

PRODUCT AXIOM. *To any two elements a' and a'' in $\mathfrak{A}$ there corresponds a unique third $a = a' \cdot a''$.*

Hence groupoids may be said shortly to be *systems with one operation.* The product axiom may be expressed also in the symbolic form

$$\mathfrak{A} \cdot \mathfrak{A} = \mathfrak{A}^2 \leqq \mathfrak{A}. \tag{1}$$

The *order* of a groupoid is the number of its elements. A subset $\mathfrak{B}$ of $\mathfrak{A}$ shall be called a *sub-groupoid* if

$$\mathfrak{B}^2 \leqq \mathfrak{B}. \tag{2}$$

Any set of elements of

$$a_1, a_2, \cdots$$

will generate a sub-groupoid

$$\mathfrak{A}' = \{a_1, a_2, \cdots\}$$

consisting of all possible products with a finite number of factors of the a_i. Since the associative law usually does not hold one has to distinguish the various ways of combining the factors. The simplest case is the *cyclic groupoid*

$$\mathfrak{C} = \{a\}$$

consisting of all possible products of one element. A *power* of a is any element belonging to $\{a\}$.

The *union* $[\mathfrak{B}, \mathfrak{C}]$ of two sub-groupoids $\mathfrak{B}$ and $\mathfrak{C}$ is the groupoid generated by the elements of $\mathfrak{B}$ and $\mathfrak{C}$ and the *cross-cut* $(\mathfrak{B}, \mathfrak{C})$ is the groupoid of their common elements. It is easily seen that the union and cross-cut as here defined satisfy the ordinary structure axioms. If we agree to consider the void set as a sub-groupoid if there exists sub-groupoids without common elements we have:

THEOREM 1. *The sub-groupoids of a given groupoid form a structure.*

Usually we shall have to impose further axiomatic restrictions on the groupoids considered. We shall say that $\mathfrak{A}$ is a *homogeneous groupoid* if it satisfies the axiom:

HOMOGENEITY AXIOM. *Each element of $\mathfrak{A}$ is a product.*

This means that for each a in $\mathfrak{A}$ we can determine other elements x and y such that

$$a = x \cdot y.$$

This again may be expressed as

$$\mathfrak{A}^2 = \mathfrak{A} \tag{3}$$

which is a stronger form of (1).

When $\mathfrak{A}$ is a non-homogeneous groupoid it may be decomposed into disjoint classes by the following method. A product

$$a = a_1 \cdot a_2 \cdot \cdots a_n \tag{4}$$

shall be said to have the *length* n. Let us again recall that since the associative law does not hold the product (4) is not uniquely defined but depends on the way in which the factors are combined. We define $\mathfrak{A}^n$ as the set of all elements of $\mathfrak{A}$ representable as products of length n. Then $\mathfrak{A}^n$ is seen to be a sub-groupoid and we have

$$\mathfrak{A} \geqq \mathfrak{A}^2 \geqq \mathfrak{A}^3 \geqq \cdots.$$

The elements of $\mathfrak{A}$ which are not products belong to

$$\Delta_1(\mathfrak{A}) = \mathfrak{A} - \mathfrak{A}^2$$

and in general those elements which are representable as products of length n but not of length $n+1$ belong to

$$\Delta_n(\mathfrak{A}) = \mathfrak{A}^n - \mathfrak{A}^{n+1}.$$

This gives us the decomposition of the elements of $\mathfrak{A}$ into the disjoint classes

$$\mathfrak{A} = \Delta_1(\mathfrak{A}) + \Delta_2(\mathfrak{A}) + \cdots.$$

Obviously those elements of $\mathfrak{A}$ which are representable as products of arbitrary length belong to all $\mathfrak{A}^n$. Hence we may say:

THEOREM 2. *The cross-cut*

$$\mathfrak{D} = (\mathfrak{A}, \mathfrak{A}^2, \cdots)$$

is the maximal homogeneous sub-groupoid of the given groupoid $\mathfrak{A}$.

It is clear that any other homogeneous sub-groupoid must be contained in $\mathfrak{D}$. If for some n we have $\mathfrak{A}^n = \mathfrak{A}^{n+1}$ then $\mathfrak{D} = \mathfrak{A}^n$ and we may call n the *order of non-homogeneity of* $\mathfrak{A}$.

If $\mathfrak{B}$ and $\mathfrak{C}$ are sub-groupoids of $\mathfrak{A}$, let us denote by $\mathfrak{B} \cdot \mathfrak{C}$ the totality of elements $b \cdot c$. Those sub-groupoids $\mathfrak{B}$ of $\mathfrak{A}$ for which

$$\mathfrak{A} \cdot \mathfrak{B} \leqq \mathfrak{B} \qquad (\mathfrak{B} \cdot \mathfrak{A} \leqq \mathfrak{B})$$

shall be called *left-ideal* (right-ideal) sub-groupoids. The study of these

groupoids would lead us to a multiplication theory of ideals and rings. In the present paper we are however more interested in systems approaching groups where such ideal groupoids do not exist.

The next considerations lead us to systems in which the following condition is fulfilled:

EXISTENCE OF QUOTIENTS. *To each pair a and b in $\mathfrak{A}$ there exist at least one x and one y such that*

$$a \cdot x = b, \qquad y \cdot a = b.$$

This condition may also be expressed

$$a \cdot \mathfrak{A} = \mathfrak{A} \cdot a = \mathfrak{A} \tag{5}$$

for every a. The solutions of the equations (5) are unique if we impose further:

CANCELLATION LAW. *When*

$$ax = ay \quad \text{or} \quad xa = ya \tag{6}$$

then $x = y$.

The two preceding conditions together may then be formulated as a single axiom:

QUOTIENT AXIOM. *To each pair a and b there exists a unique x and y such that*

$$ax = b, \qquad ya = b. \tag{7}$$

A groupoid satisfying the quotient axiom shall be called a *quasi-group*. Quasi-groups differ from the ordinary groups in the properties that no unit element, hence no inverse need exist and the associative law is usually not satisfied. For finite groupoids the cancellation law and the quotient axiom are equivalent since they both imply that each element of $\mathfrak{A}$ is contained in every column and line of the corresponding Cayley square. If one only supposes that one of the equations (7) are uniquely solvable we may call the corresponding system a left-hand or right-hand quasi-group.

2. Co-set expansions. Since a great part of the theory of groups depends upon the theory of co-set expansions it seems natural to investigate when a co-set expansion can be obtained in a groupoid. Let $\mathfrak{A}$ be the given groupoid and $\mathfrak{B}$ any sub-groupoid. The totality of elements $a \cdot \mathfrak{B}(\mathfrak{B} \cdot a)$ shall be called a left (right) *co-set* of $\mathfrak{A}$ with respect to $\mathfrak{B}$. We define:

CO-SET EXPANSION. *We shall say that there exists a left co-set expansion of* $\mathfrak{A}$ *with respect to* $\mathfrak{B}$ *when the co-sets of* $\mathfrak{A}$ *with respect to* $\mathfrak{B}$ *form a system of disjoint sets exhausting* $\mathfrak{A}$. *Furthermore no co-set shall contain equal elements.*

These conditions may be expressed

I. The equation

$$x \cdot b = a$$

has a solution for all a and some b in $\mathfrak{B}$.

II. Any relation

$$a_1 \cdot b_1 = a_2 \cdot b_2, \qquad b_1, b_2 \text{ in } \mathfrak{B}$$

implies

$$a_1\mathfrak{B} = a_2\mathfrak{B}.$$

III. Any relation

$$ab_1 = ab_2$$

implies $b_1 = b_2$.

Condition I implies namely that every element of $\mathfrak{A}$ belongs to some co-set while II shows that two co-sets are either identical or disjoint and III shows that no co-set contains equal elements.

From I follows immediately:

THEOREM 3. *Co-set expansions can only exist in homogeneous groupoids.*

Condition III permits us to draw the ordinary conclusion that if $\mathfrak{A}$ and $\mathfrak{B}$ have the finite orders $N_{\mathfrak{A}}$ and $N_{\mathfrak{B}}$ then $N_{\mathfrak{B}}$ divides $N_{\mathfrak{A}}$ and the quotient $N_{\mathfrak{A}}/N_{\mathfrak{B}}$ is equal to the number of distinct co-sets.

Let us next determine the conditions for the existence of left co-set expansions for all sub-groupoids of $\mathfrak{A}$. When one agrees to consider $\mathfrak{A}$ itself as a sub-groupoid, it can be shown:

THEOREM 4. *The necessary and sufficient condition that there shall exist a left co-set expansion for all sub-groupoids of a given groupoid* $\mathfrak{A}$ *is that:*

I. *To each* a *and* b *there shall exist such a* b_a *in the cyclic groupoid* $\{b\}$ that the equation

$$x \cdot b_a = a$$

has a solution.

II. *Any relation*

$$a_1 \cdot b_1 = a_2 \cdot b_2$$

shall imply

$$a_1\mathfrak{B} = a_2\mathfrak{B}$$

where $\mathfrak{B}$ *is any groupoid containing* b_1 *and* b_2.

III. *The left-hand cancellation law shall be satisfied.*

The proof is immediate. Let us observe that the first condition implies that every element in $\mathfrak{A}$ lies in a co-set with respect to any cyclic sub-groupoid of $\mathfrak{A}$ and conversely when it is satisfied any a lies in some co-set with respect to any sub-groupoid.

In the case where one does not consider $\mathfrak{A}$ to be a sub-groupoid the above conditions must be slightly modified.

3. The first associative law. From now on we shall suppose that $\mathfrak{A}$ is a *quasi-group*. Some of the results are however valid under somewhat more general conditions but we shall leave these more general formulations of the theorems to the reader.

We shall also suppose that $\mathfrak{A}$ is of finite order. This assumption is not essential in the sense that one can formulate the results in such a manner that they also hold in the infinite case. The theory is considerably simplified however for finite quasi-groups and the number of necessary conditions is reduced. One consequence of this condition is that every sub-groupoid $\mathfrak{H}$ is a quasi-group. We have namely for any h in $\mathfrak{H}$

$$h\mathfrak{H} = \mathfrak{H}h = \mathfrak{H}$$

because all these complexes contain the same number of elements.

From theorem 4 follows next:

THEOREM 5. *The necessary and sufficient condition that there exist left co-set expansions in a finite quasi-group* $\mathfrak{A}$ *with respect to any sub-groupoid is that the following condition be fulfilled:*

A. CONDITION OF ASSOCIATIVITY. *Any relation*

$$a_1 \cdot b_1 = a_2 \cdot b_2 \tag{8}$$

shall imply

$$a_1 \cdot \mathfrak{B} = a_2 \cdot \mathfrak{B} \tag{9}$$

where $\mathfrak{B}$ *is any groupoid containing* b_1 *and* b_2.

We shall now analyse the condition of associativity and show that it may actually be considered as a weak form of the associative law.

THEOREM 6. *The condition of associativity in a quasi-group* $\mathfrak{A}$ *is equivalent to:*

A_1. ASSOCIATIVE LAW. *Let* a *and* b *be arbitrary elements of* $\mathfrak{A}$. *For the fixed element* c_0 *let* d_0 *be determined such that*

$$(a \cdot b) \cdot c_0 = a \cdot d_0. \tag{10}$$

Then we have for any c

$$(a \cdot b) \cdot c = a \cdot d \tag{11}$$

where d *belongs to the groupoid* $\mathfrak{C} = \{c_0, d_0, c\}$ *generated by* c_0, d_0 *and* c.

Proof. Obviously the condition A implies A_1. To prove the converse let a relation (8) hold. Since we can write $a_1 = a_2 \cdot x$ it can be expressed

$$(a_2 \cdot x) \cdot b_1 = a_2 \cdot b_2$$

and hence by A_1

$$(a_2 \cdot x) \cdot c = a_1 \cdot c = a_2 \cdot d$$

where d belongs to the groupoid $\{b_1, b_2, c\}$. Hence we have for any groupoid $\mathfrak{B}$ containing b_1 and b_2

$$a_1 \cdot \mathfrak{B} \leqq a_2 \cdot \mathfrak{B}$$

and on account of the symmetry we obtain (9).

It may be of interest to observe that more special associative laws are implied in (11).

THEOREM 7. *The associative law* A_1 *implies*

$$(a \cdot b) \cdot c = a \cdot d$$

where d *belongs to the groupoid* $\mathfrak{C} = \{e_{ab}, b, c\}$ *generated by* b, c *and the element* e_{ab} *defined by*

$$(ab)\, e_{ab} = ab.$$

The proof of this theorem follows by identifying the last relation with (10).

4. The second associative law. The associative law A_1 gave the necessary and sufficient condition for the existence of left co-sets in a quasi-group. By imposing further conditions on the properties of the co-sets we are led to more special associative laws.

THEOREM 8. *The necessary and sufficient conditions that the co-sets shall have the property that any element in the co-set defines the same co-set is that the quasi-group have the following associative law:*

A_2. *For any set of elements a and b one has*

$$(a \cdot b) \cdot c = a \cdot d \tag{12}$$

where d belongs to $\{b, c\}$.

Proof. In order that

$$a \cdot \mathfrak{B} = (a \cdot b)\mathfrak{B} \tag{13}$$

for any groupoid $\mathfrak{B}$ containing b the condition (12) must obviously be satisfied and when (12) is satisfied (13) must also hold. The associative law A_2 implies the condition of associativity A (or A_1). If namely (8) holds we find from (13) for any $\mathfrak{B}$ containing b_1 and b_2

$$a_1 \cdot \mathfrak{B} = (a_1 \cdot b_1)\mathfrak{B} = (a_2 \cdot b_2)\mathfrak{B} = a_2 \cdot \mathfrak{B}$$

and (9) is proved.

When the associative law A_2 holds one can deduce several other important properties of the quasi-group.

THEOREM 9. *For any sub-groupoid $\mathfrak{B}$ containing b one has*

$$a \cdot \mathfrak{B} = (a \cdot b) \cdot \mathfrak{B} = (a \cdot \mathfrak{B}) \cdot b. \tag{14}$$

These complexes are all contained in $a \cdot \mathfrak{B}$ and contain the same number of elements. Furthermore:

THEOREM 10. *Any co-set $a \cdot \mathfrak{B}$ remains the same when multiplied by an element of $\mathfrak{B}$ on the right. The co-set contains its multiplier a.*

To prove the last property we observe that according to (14) we can determine an e_a in $\mathfrak{B}$ such that

$$a \cdot b = (a \cdot e_a) \cdot b$$

and when b is cancelled one finds

$$a = a \cdot e_a. \tag{15}$$

This proof gives another interesting property of the quasi-groups with the associative law A_2. In any quasi-group one can find a unique element e_a satisfying (15). We shall call e_a the *right unit* of a. The totality of right units generate a sub-groupoid which we shall call the *right unit groupoid* $\mathfrak{E}$. From (14) follows that in the case where A_2 holds the unit e_a is contained in any sub-groupoid of $\mathfrak{A}$.

THEOREM 11. *In quasi-groups satisfying the associative law A_2 the right-unit groupoid is a minimal groupoid contained in all other sub-groupoids.*

The existence of such a minimal groupoid not equal to a unit element is somewhat surprising. Let us illustrate its existence by the example of a quasi-group defined by the following multiplication table:

	1	2	3	4	5	6
1	2	1	3	5	4	6
2	1	3	2	4	6	5
3	3	2	1	6	5	4
4	5	4	6	2	1	3
5	4	6	5	1	3	2
6	6	5	4	3	2	1

This groupoid is cyclic. Its minimal groupoid consists of the elements 1, 2, 3.

THEOREM 12. *In a quasi-group satisfying the associative law A_2 let a and b be given elements. Then there exists a power b_1 of b such that*

$$(a \cdot b) \cdot b_1 = a. \tag{16}$$

The proof of this theorem follows from theorem 9 and 10 when applied to $\mathfrak{B} = \{b\}$. The relation (16) represents a weak form for *inverses.*

Among the other properties of the right unit elements in such systems let us mention

$$(a \cdot b) \cdot e_c = a \cdot b_1, \qquad (a \cdot e_b) \cdot c = a \cdot c_1$$

where b_1 and c_1 are powers of b and c respectively. If a quasi-group satisfying A_2 contains an *idempotent* element

$$a^2 = a$$

then a is an *absolute right unit* since the minimal groupoid consists of a single element a.

5. The third associative law. There exists other properties of the ordinary co-sets in groups which we cannot derive from A_2. In a group we have: Let $\mathfrak{B} > \mathfrak{C}$ be two subgroups of a given group $\mathfrak{A}$. The co-set decomposition of $\mathfrak{A}$ with respect to $\mathfrak{C}$ may be obtained by first decomposing $\mathfrak{A}$ with respect to $\mathfrak{B}$ and then substituting in the co-sets $a \cdot \mathfrak{B}$ the co-sets of $\mathfrak{B}$ with respect to

$\mathfrak{C}$. This property may be called the *transitivity* of the co-set decomposition. If we want the co-set expansions in quasi-groups to have the same property we must have

$$a \cdot (b \cdot \mathfrak{C}) = a_1 \cdot \mathfrak{C} \tag{17}$$

where a_1 is a fixed element independent of the particular element chosen in $\mathfrak{C}$. Since we want the transitivity in addition to the former properties of the co-set decomposition we can assume that A_2 holds. When $\mathfrak{C}$ is taken as the right unit groupoid $\mathfrak{E}$ we find from (17)

$$a(be_b) = ab = (ab)e_{ab} = a_1 \cdot e$$

where e is some element of $\mathfrak{E}$. Since $\mathfrak{E}$ is contained in all sub-groupoids we find from condition A

$$a(b \cdot \mathfrak{C}) = (ab)\mathfrak{C}.$$

This shows, when $\mathfrak{C}$ is taken as a cyclic group, that *the necessary and sufficient condition that the co-set decomposition in a quasi-group $\mathfrak{A}$ be transitive and have the property that every element of a co-set define the same co-set, is that the following associative law be satisfied:*

A_3. ASSOCIATIVE LAW. *For any three elements a, b, c*

$$a(bc) = (ab)c_1$$

where c_1 is a power of c.

In connection with the various associative laws which we have derived in the preceding one should mention the generalisations of the associative law considered by other authors. Most of these investigations differ from the present since they deal with associative laws in rings, hence with associative laws containing two operations. One of the most interesting purely multiplicative associative laws is due to *Suschkewitsch:*[1]

In the relation

$$(a \cdot b) \cdot c = a \cdot d \tag{18}$$

the element d does not depend on a.

This law is a special case of the second associative law A_2 as one sees by making a an element of the groupoid $\{b, c\}$. In this case there also exists an absolute right unit. To see this we write

$$(a \cdot e_a) \cdot c = a \cdot c$$

[1] A. Suschkewitsch, "On a generalization of the associative law," *Transactions of the American Mathematical Society*, vol. 31 (1929), pp. 204-214.

and according to (18) we have for any b

$$(b \cdot e_a) \cdot c = b \cdot c$$

or

$$b \cdot e_a = b.$$

This shows that $e = e_a = e_b = \cdots$ is a right unit. On the basis of the associative law (18) Suschkewitsch proves the theorem of *Lagrange* for quasi-groups, i. e. the order of a sub-quasi-group divides the order of the quasi-group. This is of course true whenever co-set decompositions exist. Suschkewitsch also considers the *associative law*

$$(19) \qquad (a \cdot b) \cdot c = a \cdot (b \cdot c_1)$$

where c_1 only depends on c. This is a special case of our associative law A_3. The most interesting result in his paper is that his quasi-groups may be obtained by certain substitutions from ordinary groups.

Another system of associative laws was considered by Miss Moufang.[2] She considers quasi-groups in which there exists a two-sided absolute unit e, for every element a there exists an inverse a^{-1}

$$a^{-1} \cdot a = a \cdot a^{-1} = e$$

with the further property that

$$a \cdot (a^{-1} \cdot b) = b, \qquad (b \cdot a^{-1}) \cdot a = b.$$

In addition one shall have the associative laws

$$(20) \qquad (a \cdot (c \cdot a)) \cdot b = a \cdot (c \cdot (a \cdot b))$$

$$(21) \qquad (a \cdot b) \cdot (c \cdot a) = a \cdot ((b \cdot c) \cdot a).$$

It is then shown that when the associative law (20) holds any *two* elements a and b generate an ordinary group. If both (20) and (21) hold, then any *three* elements for which $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ will also generate a group.

Another method of studying the properties of quasi-groups is to consider those elements a which satisfy one of the relations

$$(x \cdot y) \cdot a = x \cdot (y \cdot a)$$
$$(x \cdot a) \cdot y = x \cdot (a \cdot y)$$
$$(a \cdot x) \cdot y = a \cdot (x \cdot y)$$

for all x and y. Such investigations have been made by *Schönhardt.*[3]

---

[2] Ruth Moufang, "Struktur von Alternativkörpern," *Mathematische Annalen*, vol. 110 (1935), pp. 416-430.

[3] E. Schönhardt, "Über Lateinische Quadrate und Unionen," *Journal für Mathematik*, vol. 163 (1930), pp. 183-230.

Chapter II. Normal Quasi-Groups.

1. Definitions and associative law. Since the structural properties of ordinary groups are closely connected with the properties of normal sub-groups, we shall consider the possibility of defining normal sub-groupoids of a quasi-group. It is natural to define the *transform* $b^{(a)}$ of an element b with respect to a as the solution of the equation

$$a \cdot b = b^{(a)} \cdot a.$$

A *normal sub-groupoid* $\mathfrak{N}$ of $\mathfrak{A}$ shall then be defined as a groupoid containing the transforms of all its elements. This is equivalent to

$$a \cdot \mathfrak{N} = \mathfrak{N} \cdot a \tag{1}$$

or

$$a \cdot n = n' \cdot a \tag{2}$$

for all n and n' in $\mathfrak{N}$ and a arbitrary in $\mathfrak{A}$.

THEOREM 1. *The cross-cut* $(\mathfrak{M}, \mathfrak{N})$ *of two normal sub-groupoids is normal or vacuous.*

For an element d in the cross-cut one has

$$a \cdot d = n \cdot a = m \cdot a$$

and hence $n = m = d'$ is also in the cross-cut. Through a similar argument one finds:

THEOREM 2. *If* $\mathfrak{N}$ *is normal and* $\mathfrak{B}$ *any sub-groupoid then the cross-cut* $(\mathfrak{B}, \mathfrak{N})$ *is vacuous or normal in* $\mathfrak{B}$.

To obtain further properties of the normal sub-groupoids it is necessary to introduce some form of the associative law. We shall begin by assuming only:

N_1. ASSOCIATIVE LAW. *For any four elements we have*

$$(ab)(cd) = (\bar{a}b_1)\bar{d} = \bar{\bar{a}}(b_2\bar{\bar{d}}) \tag{3}$$

where $\bar{a}$ *and* $\bar{\bar{a}}$ *are powers of* a, $\bar{d}$ *and* $\bar{\bar{d}}$ *powers of* d *and* b_1 *and* b_2 *belong to the groupoid* $\{b, c\}$.

By making d an element of $\{b, c\}$ we find as a special case of N_1

$$(ab)c = \bar{a} \cdot b_1 \tag{4}$$

where $\bar{a}$ is a power of a and b_1 belongs to $\{b, c\}$. Similarly one finds

$$a(bc) = a_1 \cdot \bar{c} \tag{5}$$

where $\bar{c}$ is a power of c and a_1 belongs to $\{a, b\}$.

These laws are weaker than A_2 and they are not sufficient to obtain co-set expansions. They are however already sufficient to prove several facts about normal sub-groupoids.

We shall say that two sub-groupoids $\mathfrak{B}$ and $\mathfrak{C}$ are *permutable* if one always has for their elements

$$b \cdot c = c' \cdot b'.$$

Hence a normal sub-groupoid is permutable with every other groupoid.

THEOREM 3. *Let $\mathfrak{B}$ and $\mathfrak{C}$ be two permutable groupoids. Then the elements of their union*

$$\mathfrak{U} = [\mathfrak{B}, \mathfrak{C}] \tag{6}$$

have one of the forms

$$u = b, \quad u = c, \quad u = b \cdot c. \tag{7}$$

Proof. The union $\mathfrak{U}$ consists of all elments of $\mathfrak{B}$ and $\mathfrak{C}$ and elements obtained by repeated products of these. Hence the theorem may be deduced by induction from the following relations which are all consequences of the associative laws (3), (4) and (5):

$$b_1 \cdot (b \cdot c) = b_2 \cdot \bar{c}$$
$$c_1 \cdot (b \cdot c) = c_1 \cdot (c_2 \cdot b_2) = c_3 \cdot \bar{b}_2$$
$$(b_1 c_1)(bc) = (c_2 b_2)(bc) = \bar{c}_2 (b_3 \bar{c}) = \bar{c}_2 (c_4 b_4) = c_5 b_5.$$

As a corollary it follows from theorem 3 that if $\mathfrak{B}$ is normal and $\mathfrak{C}$ arbitrary the union (6) will consist of elements of the form (7).

A fundamental result is the following:

THEOREM 4. *Let $\mathfrak{B}$ and $\mathfrak{C}$ be permutable groupoids and $\mathfrak{D}$ some groupoid containing $\mathfrak{B}$. Then the three groupoids satisfy the Dedekind relation*

$$(\mathfrak{D}, [\mathfrak{B}, \mathfrak{C}]) = [\mathfrak{B}, (\mathfrak{D}, \mathfrak{C})]. \tag{8}$$

Proof. Obviously the right-hand side of (8) is contained in the left.

To show the converse we observe that each element of $[\mathfrak{B}, \mathfrak{C}]$ has one of the forms (7). Hence the left-hand cross-cut consists of elements satisfying one of the relations

$$d = b, \quad d = c, \quad d = bc.$$

In the two first cases the element is contained in $\mathfrak{B}$ or $(\mathfrak{D}, \mathfrak{C})$. In the last case b belongs to $\mathfrak{D}$ since $\mathfrak{D} > \mathfrak{B}$, and since $\mathfrak{D}$ is a quasi-group the element c must belong to $\mathfrak{D}$, hence to $(\mathfrak{D}, \mathfrak{C})$. Finally (8) also holds when one of the cross-cuts is void.

2. The law of isomorphism. Let $\mathfrak{B} > \mathfrak{C}$ be two sub-groupoids of $\mathfrak{A}$. With each such pair we associate a *quotient structure* consisting of all sub-groupoids $\bar{\mathfrak{C}}$ of $\mathfrak{A}$ such that

$$\mathfrak{B} \geqq \bar{\mathfrak{C}} \geqq \mathfrak{C}.$$

Furthermore we shall say that two quotients are *structure isomorphic*

$$\mathfrak{B}/\mathfrak{C} \sim \mathfrak{B}_1/\mathfrak{C}_1$$

when there exists a one-to-one correspondence between their sub-groupoids $\bar{\mathfrak{C}} \rightleftarrows \bar{\mathfrak{C}}_1$ preserving union and cross-cut. We can then prove:

THEOREM 5. *Let $\mathfrak{B}$ and $\mathfrak{C}$ be groupoids such that $\mathfrak{C}$ is normal in the union $[\mathfrak{B}, \mathfrak{C}] = \mathfrak{U}$ and the cross-cut $(\mathfrak{B}, \mathfrak{C}) = \mathfrak{D}$ is not void. Then $\mathfrak{D}$ is normal in $\mathfrak{B}$ and there exists a structure isomorphism*

$$[\mathfrak{B}, \mathfrak{C}]/\mathfrak{C} \sim \mathfrak{B}/(\mathfrak{B}, \mathfrak{C}). \tag{9}$$

Proof. From theorem 2 follows that $\mathfrak{D}$ is normal in $\mathfrak{B}$. Next let $\bar{\mathfrak{D}}$ and $\bar{\mathfrak{C}}$ be any groupoids such that

$$\mathfrak{B} \geqq \bar{\mathfrak{D}} \geqq \mathfrak{D}, \qquad [\mathfrak{B}, \mathfrak{C}] \geqq \bar{\mathfrak{C}} \geqq \mathfrak{C}. \tag{10}$$

We shall then show that the correspondence

$$\bar{\mathfrak{D}} \to [\mathfrak{C}, \bar{\mathfrak{D}}], \qquad \bar{\mathfrak{C}} \to (\mathfrak{B}, \bar{\mathfrak{C}}) \tag{11}$$

defines the structure isomorphism (9). One finds first from the Dedekind relation (8) that

$$\bar{\mathfrak{D}} = (\mathfrak{B}, [\mathfrak{C}, \bar{\mathfrak{D}}]), \qquad \bar{\mathfrak{C}} = [\mathfrak{C}, (\mathfrak{B}, \bar{\mathfrak{C}})]$$

so that (11) gives a one-to-one correspondence between the two structures. From these relations one obtains

$$[\bar{\mathfrak{D}}_1, \bar{\mathfrak{D}}_2] \to [\mathfrak{C}, \bar{\mathfrak{D}}_1, \bar{\mathfrak{D}}_2] = [\bar{\mathfrak{C}}_1, \bar{\mathfrak{C}}_2]$$
$$(\bar{\mathfrak{D}}_1, \bar{\mathfrak{D}}_2) \to [\mathfrak{C}, (\bar{\mathfrak{D}}_1, \bar{\mathfrak{D}}_2)] = [\mathfrak{C}, (\mathfrak{B}, [\bar{\mathfrak{C}}_1, \bar{\mathfrak{C}}_2])] = [\bar{\mathfrak{C}}_1, \bar{\mathfrak{C}}_2]$$

completing the proof.

Some other properties of these normal sub-groupoids may be derived. One finds that their properties correspond exactly to those of the semi-normal elements introduced by *Ore*[4] for an arbitrary structure. One can derive the same weak theorems for chains of normal sub-groupoids as for chains of semi-normal elements in structures. These properties are however not strong enough to derive a theorem of Jordan-Hölder, even in the weak form that two normal chains have the same length.

From now we shall consider quasi-groups with both right and left co-set expansions. Hence we shall assume as in chap. I:

A_2. ASSOCIATIVE LAW. *For any three elements*

$$(ab)c = a \cdot d, \qquad a(bc) = f \cdot c$$

where d belongs to $\{b, c\}$ *and f to* $\{a, b\}$.

This assumption has various consequences. We have seen that it implies the existence of a minimal sub-groupoid $\mathfrak{E}$ contained in all other sub-groupoids and $\mathfrak{E}$ is generated by all right and left units of the elements of $\mathfrak{A}$. In this case the cross-cut of two groupoids cannot be void. Theorem 3 takes the simpler form: *If* $\mathfrak{B}$ *and* $\mathfrak{C}$ *are permutable then the union* $[\mathfrak{B}, \mathfrak{C}]$ *consists* of the elements $b \cdot c$.

By making $b = e_a$ in (3) one finds that N_1 now implies:

A_4. ASSOCIATIVE LAW. *For any three elements*

$$(ab)c = \bar{a}(\bar{b}\bar{c})$$

where $\bar{a}$, $\bar{b}$, $\bar{c}$ *are powers of a, b, c respectively.*

In order to combine A_2 and A_4 we are led to:

A_5. *For any three elements a relation*

(12) $$(ab)c = a(\bar{b}\bar{c})$$

shall hold both ways, where $\bar{b}$ *is a power of b and conversely,* $\bar{c}$ *a power of c and conversely. Similarly we assume*

[4] Oystein Ore, "On the theorem of Jordan-Hölder," *Transactions of the American Mathematical Society*, vol. 41 (1937), pp. 266-275.

$$a_1(b_1c_1) = (\bar{a}_1 \cdot \bar{b}_1) \cdot c_1. \tag{13}$$

This law shall be assumed to hold in the following. Obviously it implies A_2 and A_4. It also implies N_1 as one sees from the relations

$$(ab)(cd) = ((ab)\bar{c}) \cdot \bar{d} = (a(\bar{b}\bar{c})) \cdot \bar{d}.$$

Let us now return to theorem 5. We expand $\mathfrak{B}$ in co-sets with regard to $\mathfrak{D}$

$$\mathfrak{B} = \mathfrak{D} + b_2\mathfrak{D} + \cdots + b_k\mathfrak{D}. \tag{14}$$

We can then show that

$$\mathfrak{U} = \mathfrak{C} + b_2\mathfrak{C} + \cdots + b_k\mathfrak{C} \tag{15}$$

is a co-set expansion of $\mathfrak{U}$. First no two co-sets (15) can be equal because it would lead to the equality of two co-sets (14). Secondly every element $b \cdot c$ of $\mathfrak{U}$ belongs to a co-set, because if one writes $b = b_i \cdot d$ one finds

$$b \cdot c = (b_i \cdot d)c = b_i \cdot c_1.$$

Thus we have defined a one-to-one correspondence

$$b_i\mathfrak{D} \rightleftarrows b_i\mathfrak{C} \tag{16}$$

between the co-sets of the two quotients (9). This correspondence is seen to define the structure isomorphism (11) between the two quotients because if

$$\bar{\mathfrak{D}} = \mathfrak{D} + \bar{b}_2\mathfrak{D} + \cdots + \bar{b}_t\mathfrak{D}$$

then one finds that the corresponding $\bar{\mathfrak{C}}$ in (11) is given by

$$\bar{\mathfrak{C}} = \mathfrak{C} + \bar{b}_2\mathfrak{C} + \cdots + \bar{b}_t\mathfrak{C}.$$

This shows furthermore that the indices of corresponding quotients $\bar{\mathfrak{D}}/\mathfrak{D}$ and $\bar{\mathfrak{C}}/\mathfrak{C}$ are the same. Such a correspondence of structures may be called a *strong* structure isomorphism ($\mathfrak{A} \simeq \mathfrak{A}'$).

THEOREM 6. *Let $\mathfrak{C}$ be normal in $[\mathfrak{B}, \mathfrak{C}]$. Then $(\mathfrak{B}, \mathfrak{C})$ is normal in $\mathfrak{B}$ and there exists a strong structure isomorphism*

$$[\mathfrak{B}, \mathfrak{C}]/\mathfrak{C} \simeq \mathfrak{B}/(\mathfrak{B}, \mathfrak{C})$$

defined by a one-to-one correspondence between the co-sets of the two quotients.

3. Decomposition theorems. We shall continue to assume the associative law A_5. We prove first:

THEOREM 7. *The union of two normal sub-groupoids is again normal.*

Proof. The union $[\mathfrak{N}, \mathfrak{M}]$ consists of elements of the form $n \cdot m$. The relation (2) is obtained by the following reduction

$$a(nm) = (a\bar{n})\bar{m} = m_1 \cdot (a\bar{n}) = m_1(n_1 \cdot a) = \bar{m}_1 \cdot \bar{n}_1) \cdot a.$$

By combining this with theorem 1 and theorem 4 we obtain the important result:

THEOREM 8. *When the associative law A_5 holds the normal sub-groupoids of $\mathfrak{A}$ form a Dedekind structure.*

From the theory of Dedekind structures [5] now follows that the ordinary theorems of decomposition must hold for our quasi-groups.

Let us call a chain of sub-groupoids $\mathfrak{B}_1 > \mathfrak{B}_2 > \cdots$ a *principal chain* when each term is normal in the first. It may be called a *maximal principal chain* when no further normal sub-groupoid of $\mathfrak{B}_1$ can be intercalated in it. From the theory of structures and the preceding results, follows:

THEOREM 9. *Let*

$$\begin{gathered} \mathfrak{B}_0 > \mathfrak{B}_1 > \cdots > \mathfrak{B}_r \\ \mathfrak{B}_0 = \mathfrak{C}_0 > \mathfrak{C}_1 > \cdots > \mathfrak{C}_s = \mathfrak{B}_r \end{gathered} \tag{17}$$

be two principal chains connecting $\mathfrak{B}_0$ and $\mathfrak{B}_r$. These chains may be refined into new principal chains consisting of the terms

$$\mathfrak{B}_{i,j} = [\mathfrak{B}_i, (\mathfrak{C}_j, \mathfrak{B}_{i-1})], \qquad \mathfrak{C}_{k,l} = [\mathfrak{C}_k, (\mathfrak{B}_l, \mathfrak{C}_{k-1})] \tag{18}$$

The two new principal chains have the same length and the quotient structures defined by consecutive terms are strongly structure isomorphic in some order.

Since we have only considered finite quasi-groups this theorem is equivalent to the theorem of Jordan-Hölder.

THEOREM 10. *Any two maximal principal chains connecting two normal sub-groupoids of $\mathfrak{A}$ have the same number of terms and the corresponding quotient structures are strongly structure isomorphic in some order.*

Among the various other decomposition theorems for quasi-groups which are consequences of the theory of Dedekind structures let us only mention the analogue of the theorem of *Schmidt-Remak*. We shall call a representation

[5] Oystein Ore, "On the foundation of abstract algebra. I and II," *Annals of Mathematics*, vol. 36 (1935), pp. 406-437 and vol. 37 (1936), pp. 265-292.

$$\mathfrak{A} = [\mathfrak{B}, \mathfrak{C}], \qquad (\mathfrak{B}, \mathfrak{C}) = \mathfrak{D}$$

a *direct decomposition* of $\mathfrak{A}$ over $\mathfrak{D}$. This concept may be extended to any number of components.

THEOREM 11. *Let*

$$\mathfrak{A} = [\mathfrak{B}_1, \mathfrak{B}_2, \cdots, \mathfrak{B}_r] = [\mathfrak{C}_1, \cdots, \mathfrak{C}_s] \tag{19}$$

be two direct normal decompositions of $\mathfrak{A}$ into direct indecomposable sub-groupoids over a groupoid $\mathfrak{D}$. Then the number of components is the same in both representations and the quotients $\mathfrak{B}_i/\mathfrak{D}$ and $\mathfrak{C}_j/\mathfrak{D}$ are strongly structure isomorphic in some order. Any component in one decomposition may be replaced by a suitable component of the other to give a new direct decomposition.

If we call a chain $\mathfrak{B}_1 > \mathfrak{B}_2 > \cdots$ a *normal chain* when each term is normal in the preceding, then one can also prove that the theorems 9 and 10 hold for normal chains, assuming as before that the associative law A_5 holds. The proofs are somewhat more complicated however and depend on certain preliminary results which we shall now deduce.

THEOREM 12. *In the correspondence of theorem* 6

$$[\mathfrak{B}, \mathfrak{C}]/\mathfrak{C} \simeq \mathfrak{B}/(\mathfrak{B}, \mathfrak{C})$$

a normal sub-groupoid in one structure corresponds to a normal sub-groupoid in the other.

Proof. Let $\bar{\mathfrak{C}}$ in (10) be a normal sub-groupoid of $[\mathfrak{B}, \mathfrak{C}]$ containing $\mathfrak{C}$. Then $\bar{\mathfrak{C}}$ consists of elements $\bar{b} \cdot c$ where the $\bar{b}$ form a sub-groupoid $\bar{\mathfrak{B}}$ of $\mathfrak{B}$ corresponding to $\bar{\mathfrak{C}}$ in (10). Obviously we then have for any element b of $\mathfrak{B}$

$$b \cdot \bar{b} = \bar{b}_1 \cdot b$$

where $\bar{b}_1$ must belong to $\mathfrak{B}$ and $\bar{\mathfrak{C}}$, hence to $\bar{\mathfrak{B}} = (\mathfrak{B}, \bar{\mathfrak{C}})$. This shows that $\bar{\mathfrak{B}}$ is normal in $\mathfrak{B}$. To prove conversely that if $\bar{\mathfrak{B}}$ is normal in $\mathfrak{B}$ then the corresponding $\bar{\mathfrak{C}}$ is normal in $[\mathfrak{B}, \mathfrak{C}]$ is somewhat more complicated. The elements of $\bar{\mathfrak{C}}$ are of the form $\bar{b} \cdot c$ and we shall have to prove a relation

$$(b_1 \cdot c_1) \cdot (\bar{b} \cdot c) = (\bar{b}_2 \cdot c_2) \cdot (b_1 \cdot c_1). \tag{20}$$

We obtain successively from A_5 and the normality conditions

$$(b_1 \cdot c_1) \cdot (\bar{b} \cdot c) = ((b_1 c_1)\bar{b}_2)c_2 = (b_1 \cdot (c_3 \cdot \bar{b}_3))c_2 = (b_1(\bar{b}_3 \cdot c_4)) \cdot c_2$$
$$= ((b_1\bar{b}_4)c_5)c_2 = (b_1\bar{b}_4) \cdot c_6 = (\bar{b}_5 \cdot b_1)c_6 = c_7(\bar{b}_5 \cdot b_1) = (c_8 \cdot \bar{b}_6) \cdot b_1.$$

Hence one always has a relation

$$(b_1 \cdot c_1) \cdot \bar{c} = \bar{c}_1 \cdot b_1. \tag{21}$$

According to theorem 12, chap. I we can write

$$b_1 = (b_1 \cdot c_1) \cdot c_9 = c_{10} \cdot (b_1 \cdot c_1)$$

and when this is substituted in (21) a relation of the form (20) is obtained.

The second result which we shall need is the analogue of a lemma of *Zassenhaus* in ordinary group theory.

THEOREM 13. *Let* $\bar{\mathfrak{B}}$ *be normal in* $\mathfrak{B}$ *and* $\bar{\mathfrak{C}}$ *normal in* $\mathfrak{C}$. *Then*

$$\bar{\mathfrak{M}} = [\bar{\mathfrak{B}}, (\mathfrak{B}, \bar{\mathfrak{C}})], \qquad \bar{\mathfrak{N}} = [\bar{\mathfrak{C}}, (\mathfrak{C}, \bar{\mathfrak{B}})]$$

are normal in

$$\mathfrak{M} = [\bar{\mathfrak{B}}, (\mathfrak{B}, \mathfrak{C})], \qquad \mathfrak{N} = [\bar{\mathfrak{C}}, (\mathfrak{B}, \mathfrak{C})]$$

respectively and there exists a strong structure isomorphism

$$\mathfrak{M}/\bar{\mathfrak{M}} \simeq \mathfrak{N}/\bar{\mathfrak{N}}. \tag{22}$$

Proof. To show that $\bar{\mathfrak{M}}$ is normal in $\mathfrak{M}$ we observe that every element of $\mathfrak{M}$ has the form $\bar{b} \cdot d$ where $\bar{b}$ belongs to $\bar{\mathfrak{B}}$ and d to $(\mathfrak{B}, \mathfrak{C})$, while every element of $\bar{\mathfrak{M}}$ has the form $\bar{b} \cdot \bar{d}$ where $\bar{d}$ belongs to $(\mathfrak{B}, \bar{\mathfrak{C}})$. We shall have to prove a relation

$$(\bar{b} \cdot d)(\bar{b}_1 \bar{d}_1) = (\bar{b}_2 \cdot \bar{d}_2)(\bar{b} \cdot d)$$

and this follows in about the same manner as the relation (20). To prove (22) we write $\mathfrak{D} = (\mathfrak{B}, \mathfrak{C})$ and find from theorem 4

$$[\bar{\mathfrak{M}}, \mathfrak{D}] = \mathfrak{M}$$
$$(\bar{\mathfrak{M}}, \mathfrak{D}) = [(\mathfrak{B}, \bar{\mathfrak{C}}), (\bar{\mathfrak{B}}, \mathfrak{C})].$$

According to theorem 6 we then obtain

$$\mathfrak{M}/\bar{\mathfrak{M}} \simeq (\mathfrak{B}, \mathfrak{C})/[(\mathfrak{B}, \bar{\mathfrak{C}}), (\bar{\mathfrak{B}}, \mathfrak{C})]$$

and the same result is found for $\mathfrak{N}/\bar{\mathfrak{N}}$.

We have now:

THEOREM 14. *The refinement theorem* 9 *and the theorem of Jordan-Hölder* 10 *hold also for normal chains.*

The proof follows by making the same refinements (18) and applying theorem 13.

4. Further problems. It might now be natural to ask when a quotient groupoid exists. This leads us to quasi-groups satisfying

A_3. ASSOCIATIVE LAW. *For any three elements*

$$(ab)c = a(b\bar{c}), \qquad a(bc) = (\bar{a}b)c \tag{23}$$

where $\bar{a}$ and $\bar{c}$ are powers of a and c respectively.

This law implies all the preceding associative laws. We have obtained it formerly as a condition for the transitivity of co-set expansions. To show that it implies the existence of a quotient groupoid let us expand $\mathfrak{A}$ with respect to a normal sub-groupoid

$$\mathfrak{A} = \mathfrak{N} + a_2\mathfrak{N} + \cdots + a_t\mathfrak{N}.$$

For the product of two co-sets we find

$$\begin{aligned}(a_i\mathfrak{N})(a_j \cdot \mathfrak{N}) &= ((a_i \cdot \mathfrak{N})a_j) \cdot \mathfrak{N} = ((\mathfrak{N}a_i)a_j) \cdot \mathfrak{N} \\ &= (\mathfrak{N}(a_ia_j))\mathfrak{N} = ((a_ia_j)\mathfrak{N})\mathfrak{N} = (a_ia_j)\mathfrak{N}.\end{aligned}$$

This shows that the co-sets form a quasi-group $\mathfrak{A}/\mathfrak{N}$ satisfying the same associative law as $\mathfrak{A}$. There exists a unit element in $\mathfrak{A}/\mathfrak{N}$, namely $\mathfrak{N}$. In this case we can strengthen theorem 6 to an isomorphism between the quotient groupoids

$$[\mathfrak{B}, \mathfrak{C}]/\mathfrak{C} \cong \mathfrak{B}/(\mathfrak{B}, \mathfrak{C}).$$

The quotient groupoid $\mathfrak{A}/\mathfrak{N}$ obviously gives a homomorphic image of $\mathfrak{A}$. One can show conversely:

THEOREM 15. *All quasi-groups with unit element homomorphic to $\mathfrak{A}$ are isomorphic to a quotient groupoid $\mathfrak{A}/\mathfrak{N}$. When $\mathfrak{A}$ has a unit element all its homomorphic images are obtained this way.*

One finds as usual that $\mathfrak{N}$ consists of those elements of $\mathfrak{A}$ which correspond to the unit element of the image.

Among the other theories which one can develop for quasi-groups let us mention the theory of *quasi-normal sub-groupoids*. One may define a sub-groupoid $\mathfrak{N}$ to be *quasi-normal* when for all a one has

$$a \cdot \mathfrak{N} = \mathfrak{N} \cdot \bar{a}$$

where $\bar{a}$ is a power of a and conversely. These quasi-normal groupoids correspond exactly to the quasi-normal groups introduced by *Ore*[6] in ordinary groups and a similar theory may be developed. For a large part of the theory it is sufficient to assume the associative law A_4. Let us observe that this law from the point of view of the theory of structures has similar properties to those of the ordinary associative law, since by the change of parentheses the elements are left in the same groupoid.

One can also develop the theory of two-sided co-sets in quasi-groups. If $\mathfrak{H}$ and $\mathfrak{K}$ are two sub-groupoids, then the totality of elements in the complex

$$\mathfrak{H}(a\mathfrak{K}) \tag{24}$$

is called a *two-sided co-set*. It is obvious that every element of a quasi-group belongs to some two-sided co-set (24). We shall say that $\mathfrak{A}$ possesses a *two-sided co-set expansion* with respect to $\mathfrak{H}$ and $\mathfrak{K}$ if

$$\mathfrak{A} = \mathfrak{H}(a_1\mathfrak{K}) + \mathfrak{H}(a_2\mathfrak{K}) + \cdots$$

where any co-sets

$$\mathfrak{H}(a\mathfrak{K}), \qquad \mathfrak{H}(b\mathfrak{K})$$

defined by arbitrary a and b are either identical or have no element in common.

One can now derive necessary and sufficient conditions for the existence of two-sided co-set expansions, with given properties much in the same manner as we have done in chap. I for one-sided co-sets. We shall not specify the various associative laws which are necessary for the various properties. Most of them are associative laws on four elements. One is also led to:

A_6. ASSOCIATIVE LAW. *For any three elements*

$$a(bc) = (\bar{a}b) \cdot \bar{c}$$

where $\bar{a}$ and $\bar{c}$ are powers of a and c and conversely.

This law does not imply one-sided co-set expansions. If one wishes both one- and two-sided co-set expansions one is again led to the associative law A_3.

THEOREM 16. *Let us suppose that the associative law A_3 holds. Then the quasi-group $\mathfrak{A}$ has two-sided co-set expansions with respect to any two sub-groupoids $\mathfrak{H}$ and $\mathfrak{K}$. Furthermore*

$$\mathfrak{H}(a\mathfrak{K}) = (\mathfrak{H}a)\mathfrak{K}.$$

---

[6] Oystein Ore, "Structures and group theory," *Duke Mathematical Journal*, vol. 3 (1937), pp. 149-174.

Any element in a two-sided co-set defines the same co-set and a co-set remains the same by left multiplication with an element of $\mathfrak{H}$ *or right multiplication with an element of* $\mathfrak{K}$.

Another important problem is the study of the special theory of *Abelian quasi-groups*. Equally interesting is also the question of the conditions for the validity of the *Sylow theory of groupoids*. All these problems must however be reserved for a later occasion.

Other problems of interest are the determination of all quasi-groups with given associative laws and their relation to ordinary groups. By certain operations one can derive quasi-groups from ordinary groups. A special case of such quasi-groups is given by the quasi-groups considered by *Suschkewitsch*.

YALE UNIVERSITY.

CORRECTIONS.

OTTO SZÁSZ, *On the partial sums of certain Fourier series.*

P. 705, l. 3, the formula on this line should be numbered (31);

P. 705, l. 6 from bottom, read $[K_q(t-x) - K_q(t+x)]$ instead of $[K_q(t-x) + K_q(t+x)]$;

P. 706, l. 1, read $\sum_{0<\lambda_n\leq\mu}$ instead of $\sum_{0<\lambda_n\leq\omega}$.

THE JOHNS HOPKINS PRESS · BALTIMORE

American Journal of Mathematics. Edited by E. T. Bell, E. W. Chittenden, Abraham Cohen, F. D. Murnaghan and J. F. Ritt. Quarterly. 8vo. Volume LIX in progress. $7.50 per volume. (Foreign postage, fifty cents.)

American Journal of Philology. Edited by Tenney Frank with the coöperation of H. Cherniss, K. Malone, B. D. Meritt, and D. M. Robinson. Quarterly. 8vo. Volume LVIII in progress. $5 per volume. (Foreign postage, twenty-five cents.)

Biologia Generalis. (International Journal of Biology). Founded by Leopold Löhner, Graz; Raymond Pearl, Baltimore, and Vladislav Růžička, Prague. It is now edited by O. Abel, L. Adametz, O. Porsch, C. Schwarz, J. Versluys and R. Wasicky of Vienna. 8vo. Volume XIII in progress.

Bulletin of the Institute of the History of Medicine. Edited by Henry E. Sigerist. Monthly except August and September. Volume V in progress. 8vo. Subscription $5 per year. (Foreign postage, fifty cents.)

Bulletin of the Johns Hopkins Hospital. Edited by James Bordley, III. Monthly. Volume LXI in progress. 8vo. Subscription $6 per year. (Foreign postage, fifty cents.)

Comparative Psychology Monographs. Roy M. Dorcus, Managing Editor. 8vo. Volume XIV in progress. $5 per volume.

Hesperia. Edited by William Kurrelmeyer and Kemp Malone. 8vo. Twenty-nine numbers have appeared.

Human Biology: a record of research. Raymond Pearl, Editor. Quarterly. 8vo. Volume IX in progress. $5 per volume. (Foreign postage, thirty-five cents.)

Johns Hopkins Studies in Romance Literatures and Languages. H. C. Lancaster, Editor. 8vo. Forty numbers have been published.

Johns Hopkins University Circular, including the President's Report and Catalogue of the School of Medicine. Ten times yearly. 8vo. $1 per year.

Johns Hopkins University Studies in Archaeology. David M. Robinson, Editor. 8vo. Twenty-three volumes have appeared.

Johns Hopkins University Studies in Education. Florence E. Bamberger, Editor. 8vo. Twenty-six numbers have appeared.

Johns Hopkins University Studies in Geology. Edward B. Mathews, Editor. 8vo. Eleven numbers have been published.

Johns Hopkins University Studies in Historical and Political Science. Under the direction of the Departments of History, Political Economy and Political Science. 8vo. Volume LV in progress. $5 per volume.

Modern Language Notes. Edited by H. C. Lancaster, W. Kurrelmeyer, R. D. Havens, K. Malone, H. Spencer and C. S. Singleton. Eight times yearly. 8vo. Volume LII in progress. $5 per volume. (Foreign postage, fifty cents.)

Reprint of Economic Tracts. J. H. Hollander, Editor. Fifth series in progress. Price $4.

Terrestrial Magnetism and Atmospheric Electricity. Founded by Louis A. Bauer; conducted by J. A. Fleming with the coöperation of eminent investigators. Quarterly. 8vo. Volume XLII in progress. $3.50 per volume.

Walter Hines Page School of International Relations. Five volumes have been published.

A complete list of publications will be sent upon request

Numerical Mathematical Analysis

BY

JAMES B. SCARBOROUGH

"A valuable feature of the book is the excellent collection of examples at the end of each chapter. . . . The book has many admirable features. The explanations and derivations of formulae are given in detail. . . . The author has avoided introducing new and complicated notations which, although they may conduce to brevity, are a serious stumbling block to the reader. The typography and paper are excellent."

—American Mathematics Monthly.

430 pages, 25 figures, crown 8vo, buckram, $5.50

TABLES OF $\sqrt{1-r^2}$ AND $1-r^2$

FOR USE IN

PARTIAL CORRELATION AND IN TRIGONOMETRY

BY

JOHN RICE MINER, Sc. D.

These tables fill a want long felt by practical workers in all branches of statistics. Everyone who uses the method of correlation has wished for tables from which the probable error of a coefficient of correlation could be obtained *with accuracy*. Similar tables to this have existed on a small scale, but never before have there been available tables of $1-r^2$ and $\sqrt{1-r^2}$ to 6 places of decimals, and 4 places in the argument. Not only are these tables of great usefulness in getting the probable error of a correlation coefficient, but also they have what will perhaps be their chief value in the calculations involved in the method of partial or net correlation. It is safe to say that these tables reduce the labor involved in this widely used statistical method by at least one-half.

50 pages, 8vo, cloth, $1.50

THE JOHNS HOPKINS PRESS · BALTIMORE